

IDENTIFICATION OF A NOVEL WILMS' TUMOUR 1 BINDING PROTEIN

NATALIE ANN LITTLE

Thesis submitted for the degree of

Doctor of Philosophy

University of Edinburgh

1999



Declaration

I certify that the work presented in this thesis is my own, except where otherwise stated, and has not been submitted as a degree at this, or any other university.

Natalie A. Little

Acknowledgements

First and foremost I would like to thank Rachel Davies for agreeing to take me on as her first student. Thanks Rachel for your enormous commitment towards my PhD, from the first day in the lab, to the very end when you had to plough through my thesis! I have certainly learnt a lot about science and research and even my written English has improved - I just hope that you too have gained from being my supervisor. Thanks too to my MRC and Edinburgh University supervisors Nick Hastie and Richard Meehan.

I would also like to thank C3 as a whole, for helping me out whenever it was required and for being a great lab to work in. In particular I would like to thank Mike for help with the construction of the M15 cDNA library; Laura and Carlo for allowing me to use their mouse library; Jacob for providing me with advice and materials for the whole mounts and for being a great bench partner; Colin and Aswin for Northern blots; and Adrian for help with the immunohistochemistry. I would also like to thank Mairi and Caroline from C4 for help with the GST work, and for those chats in Spanish (keep that Spanish up Caroline!); Ian for not only joining in the conversations with his 'huesos' and 'lobo' but for a good many beers and for providing the DNA for the EUCIB backcross.

Thanks too to Liz Graham for help on the *in situ* hybridisation on sections; Muriel Lee for carrying out the FISH - and succeeding even though I only provided her with cDNA; Paul Perry for fantastic help with Mac Red and teaching me how to use *SmartCapture*; Agnes for sequencing; Sandy and Norman at photography as well as Al *et al.* in computing. I am sure there are many others in the MRC who have helped me one way or another, so a general thanks to you all.

Finally I would like to thank my husband Mark, even if he did make me leave Edinburgh for Holland at my writing up stage, for help with my figures and computer troubles and for occasionally trying to show some interest in my work!

TABLE OF CONTENTS

DECLARATION.....	III
ACKNOWLEDGEMENTS.....	V
TABLE OF CONTENTS	VII
LIST OF FIGURES AND TABLES.....	XIII
ABBREVIATIONS.....	XV
ABSTRACT	XIX
CHAPTER 1: INTRODUCTION	2
1.1 CLINICAL PRESENTATION	2
1.2 HISTOLOGY	3
1.3 GENETIC LOCI ASSOCIATED WITH WT	4
1.3.1 <i>The Knudson Model And WT1</i>	4
1.3.2 <i>The WT2 Locus</i>	5
1.3.3 <i>Additional WT Genes</i>	6
1.4 CLONING THE WILMS' TUMOUR GENE, <i>WT1</i>	7
1.5 STRUCTURE OF WT1	7
1.6 FUNCTIONS OF WT1.....	9
1.6.1 <i>WT1 As A Transcription Factor</i>	9
1.6.2 <i>WT1 As A Post-Transcriptional Regulator</i>	12
1.6.3 <i>WT1 is involved in cell growth regulation and apoptosis</i>	14
1.6.4 <i>WT1 Can Interact With Several Cellular Proteins</i>	15
1.7 EXPRESSION PATTERN OF WT1	16
1.7.1 <i>The Expression Of WT1 In The Developing Kidney And The Wt1 Knockout Mouse</i>	17
1.7.2 <i>WT1 Expression in the Gonads</i>	20
1.7.3 <i>Regulation of WT1 Expression</i>	21
1.8 WT1 INVOLVED IN DISEASE.....	22
1.9 SUMMARY AND AIM OF PROJECT	24
CHAPTER 2: MATERIALS AND METHODS.....	28
2.1 BACTERIAL WORK.....	28
2.1.1 BACTERIAL STRAINS USED	28
2.1.2 <i>Media and Additives</i>	28
2.1.3 <i>Bacterial Glycerol Stocks</i>	29
2.1.4 <i>Preparing Bacterial Competent Cells</i>	29
2.1.5 <i>Transformation of Competent Cells</i>	29

2.1.6 Small Scale Preparation of Plasmid DNA	30
2.1.7 Large Scale Preparation of Plasmid DNA	30
2.2 YEAST TWO HYBRID WORK	30
2.2.1 Strain Used	30
2.2.2 Vectors	31
2.2.3 cDNA Libraries	31
2.2.4 Yeast Transformations	31
2.2.5 β -Galactosidase Yeast Filter Lifts Assay	32
2.2.6 Liquid β -Galactosidase Assay	33
2.2.7 Recovery Of Plasmids From Yeast Into <i>E. coli</i>	33
2.3 DNA MANIPULATIONS	34
2.3.1 Restriction Endonuclease Digestion of DNA	34
2.3.2 DNA Gel Electrophoresis	34
2.3.3 Dephosphorylation of Linear Plasmid DNA	34
2.3.4 Genecleaning	35
2.3.5 Ligation of DNA Molecules	35
2.3.6 Phenol:Chloroform Extraction	35
2.3.7 Ethanol Precipitation	35
2.3.8 Southern Blotting	36
2.3.9 Screening a Mouse cDNA Library	37
2.3.10 Primer Production	37
2.3.11 Polymerase Chain Reaction (PCR)	37
2.3.12 ABI Sequencing	37
2.4 RNA MANIPULATIONS	38
2.4.1 Northern Blotting	38
2.4.2 Reverse Transcriptase PCR (RT-PCR)	40
2.4.3 DIG Labelling An RNA Probe	40
2.5 IN SITU HYBRIDISATIONS	40
2.5.1 Whole Mount In Situs	40
2.5.2 In Situ Hybridisation For Sections	42
2.6 PROTEIN WORK	44
2.6.1 In Vitro Transcription and Translation of Proteins	44
2.6.2 Expressing pET6H-WT1 Proteins in Bacteria	44
2.6.3 Expressing GST-Fusion Proteins in Bacteria	45
2.6.4 Purification of the GST-Fusion Protein	46
2.6.5 BioRad Protein Assay	46
2.6.6 In Vitro Binding Assays	47
2.6.7 SDS-PAGE Electrophoresis	47

2.6.8 Detecting A Radioactive Protein.....	48
2.6.9 Western Blotting.....	48
2.7 ANTIBODY WORK.....	49
2.7.1 Antibody Production.....	49
2.7.2 Affinity Purification of NAL Antisera.....	50
2.7.3 Antibodies Used.....	50
2.8 TISSUE CULTURE.....	51
2.8.1 Cell Lines.....	51
2.8.2 Medium.....	51
2.8.3 Splitting Cells.....	51
2.8.4 Freezing Cell Lines Down.....	51
2.8.5 Growing Cell Lines Up From Frozen Stocks.....	51
2.8.6 Transient Transfections.....	52
2.8.7 β -Galactosidase Assay.....	52
2.8.8 Co-Immunoprecipitations.....	53
2.9 FLUORESCENCE.....	54
2.9.1 Immunocytochemistry.....	54
2.9.2 Immunohistochemistry.....	54
2.9.3 Fluorescent In Situ Hybridisation (FISH).....	55
CHAPTER 3: YEAST TWO HYBRID SYSTEM.....	58
3.1 INTRODUCTION TO THE YEAST TWO HYBRID SYSTEM.....	58
3.2 TESTING THE ACTIVITY OF WT1, AS WELL AS CONTROL PROTEINS, IN THE Y2H SYSTEM.....	61
3.3 SCREENING AN M15 LIBRARY.....	63
3.3.1 Generating an M15 cDNA library.....	63
3.3.2 Identifying WT1-interacting clones.....	63
3.3.3 Ensuring that the prey protein on its own does not activate the Y2H system.....	65
3.3.4 Testing the specificity for WT1 interaction.....	65
3.3.5 Summary of M15 screen.....	66
3.4 SCREENING A COMMERCIAL KIDNEY CDNA LIBRARY.....	66
3.4.1 Choice of library.....	66
3.4.2 Identifying WT1-interacting clones.....	67
3.4.3 Ensuring the clones are specific for WT1 interaction.....	67
3.4.4 Domain binding studies with the possible WT1-specific clones.....	69
3.4.5 Sequencing the clones.....	73
3.5 NAL, A CANDIDATE WT1 BINDING PROTEIN.....	74
3.5.1 NAL can bind to WT1 when expressed as the bait or the prey protein.....	74
3.5.2 NAL binds to the -KTS, zinc finger region of WT1.....	75

3.6 CONCLUSIONS	76
CHAPTER 4: IDENTIFICATION OF A NOVEL PROTEIN.....	82
4.1 SEQUENCING NAL.....	82
4.2 DATABASE MATCHES FOR NAL	82
4.2.1 NAL partially matches EST D14661	82
4.2.2 Do either NAL or the EST originate from a chimaeric clone?	82
4.2.3 The ESTs matching NAL may originate from unspliced RNA, or they could be alternative transcripts.....	86
4.3 NAL IS A NOVEL PROTEIN.....	90
4.4 ISOLATING A MOUSE CDNA HOMOLOGUE OF NAL	90
4.4.1 Screening a mouse cDNA library.....	90
4.4.2 The mouse and human clones of NAL are 94% identical.....	93
4.5 THE HUMAN NAL CLONE IS NOT QUITE FULL LENGTH	93
4.6 MAPPING NAL	94
4.6.1 Fluorescent in situ hybridisation	94
4.6.2 European collaborative interspecific mouse backcross	97
4.7 CONCLUSIONS	104
CHAPTER 5: BINDING STUDIES	108
5.1 IN VITRO BINDING ASSAYS	108
5.1.1 Problems Using Bacterially Expressed pET6HWT1.....	108
5.1.2 Problems Using Bacterially Expressed GST-WT1 Zinc Fingers	111
5.1.3 Using Bacterially Expressed GST-NAL	112
5.2 IN VIVO ASSAYS USING HA EPI TOPE-TAGGED NAL.....	117
5.2.1 Epitope Tagging NAL.....	117
5.2.2 Testing Expression Of The HA-NAL Clone.....	117
5.2.3 Binding Assays With HA-NAL.....	118
5.3 PRODUCING A POLYCLONAL ANTIBODY AGAINST NAL.....	121
5.3.1 Production Of Antibody.....	121
5.3.2 Testing The NAL Antibody	121
5.4 IN VIVO BINDING ASSAYS USING ENDOGENOUSLY EXPRESSED NAL	123
5.5 CONCLUSION	124
CHAPTER 6: EXPRESSION AND CELLULAR LOCALISATION STUDIES.....	128
6.1 DETERMINING THE EXPRESSION PATTERN USING RT-PCR	128
6.2 IMMUNOFUORESCENCE ON SECTIONS THROUGH THE KIDNEY AND TESTES	130
6.3 IN SITU HYBRIDISATION STUDIES	132
6.3.1 Whole mount in situ hybridisation	132

6.3.2 <i>In situ</i> hybridisation on sections	134
6.4 NAL IS A NUCLEAR PROTEIN.....	136
6.5 NAL CO-LOCALISES WITH SPLICING FACTORS	136
6.6 NAL REDISTRIBUTES INTO LARGE FOCI AFTER ACTINOMYCIN D TREATMENT.....	138
6.6 CONCLUSIONS.....	139
CHAPTER 7: SUMMARY AND DISCUSSION.....	144
7.1 A NOVEL PROTEIN IS IDENTIFIED WHICH BINDS WT1 IN THE YEAST TWO HYBRID SYSTEM.....	144
7.2 NAL BINDS TO WT1 BOTH <i>IN VITRO</i> AND <i>IN VIVO</i>	146
7.3 MAPPING THE MOUSE AND HUMAN NAL GENE	147
7.4 NAL IS UBIQUITOUSLY EXPRESSED AND MAY HAVE ALTERNATIVE TRANSCRIPTS	147
7.5 CELLULAR LOCALISATION AND POSSIBLE FUNCTIONS OF NAL	149
7.6 FUTURE DIRECTIONS	151
BIBLIOGRAPHY	158
APPENDIX: PRIMERS USED IN THIS THESIS	A2

List of figures and tables

Chapter 1

Figure 1.1	Schematic of the WT1 gene and protein structure	8
Table 1.1	Potential target genes regulated by WT1	11
Figure 1.2	Diagram illustrating the expression pattern of WT1 during nephrogenesis	19

Chapter 2

Table 2.1	Solutions for preparing resolving and stacking gels for Tris-glycine SDS PAGE	48
Table 2.2	List of antibodies used for Western blots and immunofluorescence	50
Table 2.3	Reagent quantities for transfection of DNA onto different sized culture plates	52

Chapter 3

Figure 3.1	Schematic illustrating the basis of the yeast two hybrid system	59
Table 3.1	Testing the activity of WT1 and control proteins in the Y2H system	62
Figure 3.2	Yeast Two Hybrid flowchart	64
Figure 3.3	β galactosidase yeast filter lift assay	67
Figure 3.4	β galactosidase assay for the specificity test	68
Figure 3.5	WT1 domain binding studies with Clone D	71
Table 3.2	WT1 domain binding studies	72
Table 3.3	NAL can bind to WT1+/+ when acting both as the prey and the bait	75
Table 3.4	NAL can bind to the zinc fingers of WT1 but only if they are -KTS	76

Chapter 4

Figure 4.1	NAL sequence showing primers used to sequence it	83
Figure 4.2	Sequence comparison between NAL and the EST, D14661	84
Figure 4.3	RT-PCR to determine whether NAL and D14661 are transcribed in cells	85
Figure 4.4	Sequence comparison between EST, h90530, and NAL	88
Figure 4.5	Model proposed for the formation of two alternative transcripts	89
Figure 4.6	RT-PCR on whole mouse RNA (E=11.5d)	91
Figure 4.7	Comparison of the peptide sequence of human NAL and mouse Nal clones	92
Figure 4.8	Sequence comparison between the end of NAL and a human EST	93
Figure 4.9	FISH analysis for NAL on human chromosome spreads	95
Figure 4.10	FISH analysis for Nal on mouse chromosome spreads	96
Figure 4.11	DdeI digests on test mice	98
Table 4.1	DdeI digest of the 3'UTR of Nal of test mice	99
Table 4.2	Haplotype data for EUCIB mice	101
Table 4.3	Mapping results for NAL using EUCIB mice	103

Chapter 5

Figure 5.1	Induction time course for pET6H-WT1	109
Figure 5.2	6H-WT1 is insoluble	110
Figure 5.3	Bacterial expression of GST- WT1 Zinc Fingers	112
Figure 5.4	GST-NAL is partially soluble	113
Figure 5.5	<i>In vitro</i> binding studies between WT1 and NAL	115
Figure 5.6	Higher stringency <i>in vitro</i> binding studies	116
Figure 5.7	Expression studies for transiently transfected HA-NAL	118
Figure 5.8	Co-immunoprecipitations of transiently transfected HA-NAL	120
Figure 5.9	Immunoblot used to test the NAL antibody	122
Figure 5.10	WT1 co-immunoprecipitates NAL	123

Chapter 6

Figure 6.1	Expression studies for NAL using RT-PCR	129
Figure 6.2	Immunofluorescence on sections through the kidney and testis	131
Figure 6.3	Whole Mount <i>In Situ</i> Hybridisation	133
Figure 6.4	<i>In situ</i> hybridisation on mouse E=10.5d cryosections	135
Figure 6.5	Immunofluorescence on M15 cells	137

Abbreviations

All units used in this thesis are Standard International (SI) units.

3-AT	3-amino-1,2,4-triazole
A	adenine
Abs	absorbance
Ac	acetate
Amp	ampicillin
AP	alkaline phosphatase
ATP	adenosine triphosphate
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
BSA	bovine serum albumin
BWS	Beckwith-Wiedemann syndrome
C	cytosine
cDNA	complementary deoxyribonucleic acid
C-terminal	carboxyl-terminal
CA	chloramphenicol
CAT	chloramphenicol acetyltransferase
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propane-sulfonate
CIP	calf intestinal phosphatase
CMV	cytomegalovirus
CTD	carboxyl terminal domain
CYH	cyclohexamide gene
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymine triphosphate
dUTP	deoxyuracil triphosphate
dsDNA	double stranded deoxyribonucleic acid
dH ₂ O	distilled water
DAM	donkey anti mouse
DAR	donkey anti rabbit
DDS	Denys-Drash syndrome
DEPC	diethyl pyrocarbonate
DIG	digoxigenin
DAPI	4',6-diamidino-2-phenylindole
DMF	dimethylformamide

DNA	deoxyribonucleic acid
DTT	dithiothreitol
E	embryonic day
EDTA	ethylenediamine tetra-acetic acid disodium salt
<i>EGFR</i>	epidermal growth factor receptor gene
<i>EGRI</i>	early growth response 1 gene
EST	expressed sequence tag
EtOH	ethanol
EUCIB	european collaborative interspecific backcross
FCS	foetal calf serum
FISH	fluorescence <i>in situ</i> hybridisation
FITC	fluorescein isothiocyanate
FS	Frasier syndrome
G	guanosine
GAM	goat anti mouse
GAR	goat anti rabbit
GDNF	glial cell line derived neurotrophic factor
GFP	green fluorescent protein
GST	glutathione S-transferase
HA	haemagglutinin tag
HGU	Human Genetics Unit
HRP	horseradish peroxidase
HSAB	high salt association buffer
Hsp70	heat shock protein 70
IGC	interchromatin granule cluster
<i>IGF1R</i>	insulin-like growth factor 1 receptor gene
<i>Igf2</i>	insulin-like growth factor 2 gene
LOH	loss of heterozygosity
LOI	loss of imprinting
ILNR	intralobar nephrogenic rest
IPTG	isopropylthio- β -D-galactosidase
LB	luria broth
leu	leucine
LMP	low melting point
LSAB	low salt association buffer
mRNA	messenger ribonucleic acid
min	minute(s)
M15	mesonephric cell line

MOPS	3-[N-morpholino]propanesulfonic acid
MRC	Medical Research Council
MSAB	medium salt association buffer
MIS	Müllerian inhibiting substance
N-terminal	amino-terminal
NAL	novel associating ligand
NBT	nitroblue tetrazolium salt
NWTS	National Wilms' Tumour Study
OD	optical density
ODC	ornithine decarboxylase
<i>Par-4</i>	prostrate apoptosis response-4 gene
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFA	paraformaldehyde
PLNR	perilobar nephrogenic rests
RACE	rapid amplification of cDNA ends
<i>RB</i>	retinoblastoma gene
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction
ssDNA	salmon sperm deoxyribonucleic acid
SAPU	Scottish Antibody Production Unit
SB	sample buffer
SDS	sodium dodecyl sulphate
<i>SF-1</i>	steroidogenic factor 1 gene
SNF1	sucrose non fermenting gene 1
SNF4	sucrose non fermenting gene 4
snRNP	small nuclear ribonucleoprotein particle
SSC	standard saline citrate
T	thymine
TBS	tris buffered saline
TBST	tris buffered saline with tween
TGS	tris/glycine/SDS buffer
TESPA	3-aminopropyl-triethoxy silane
T_m	melting temperature

Tris	tris-[hydroxymethyl]aminomethane
UAS	upstream activator sequence
UBC9	ubiquitin conjugating enzyme
UPD	uniparental paternal disomy
UTR	untranslated region
WAGR	Wilms' tumour, aniridia, genital abnormalities and mental retardation
WCE	whole cell extract
WT1	Wilms' tumour 1 gene
X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactoside
Y2H	yeast two hybrid
ZF	zinc finger

Abstract

The paediatric nephroblastoma Wilms' Tumour, affecting 1 in 10,000 children, is a disease where the disruption in the normal events of kidney development leads to tumour formation. The *WT1* gene has been shown by knockout experiments to be crucial in kidney and gonadal development. The structure of the protein, which contains four zinc fingers of the Krüppel-type, together with *in vitro* work, suggests WT1 is a transcription factor. Other work however also suggests that the protein may be involved in post-transcriptional regulation as WT1 localises and co-immunoprecipitates with splicing factors. The structure of the protein also backs up this possible role as WT1 has a putative RNA recognition motif and the zinc fingers can bind RNA *in vitro*.

The genuine functions of WT1 are thus not clear so this project set out to address what WT1 is doing at the molecular level. By looking for WT1-binding proteins, I aimed to investigate which components of the cellular machinery WT1 is interacting with, as it is only once we comprehend the molecular mechanism behind development that we will begin to understand the link with tumourigenesis.

Using a yeast two hybrid screen I identified a protein that interacts with the C-terminus of WT1. As this protein is novel, with no homologues found in the yeast or invertebrate databases, it was called NAL (Novel Associating Ligand). NAL was mapped, using FISH, to human chromosome 6q26-27, a location which possibly harbours a tumour suppressor gene. The mouse homologue was cloned and, at the amino acid level, is 94% identical to the human clone. This sequence conservation suggests that whatever the role of this novel protein may be, its structure must be important for function. Using FISH and the EUCIB resource, the mouse homologue was mapped close to the centromere on chromosome 17, a region which has conserved synteny with human 6q26-27 and which maps close to the murine T-locus.

In vitro assays showed that the interaction between NAL and WT1 could occur outside of the yeast two hybrid system and *in vivo* co-immunoprecipitation experiments revealed that the interaction is physiologically relevant as it occurs in WT1-expressing cells. Expression studies using RT-PCR and whole mount *in situ* hybridisation showed that NAL is ubiquitously expressed. The expression pattern of NAL during development of the kidney and testis was also investigated to see if, like WT1, NAL is confined to the developing nephron and Sertoli cells respectively. Immunohistochemistry revealed that NAL is expressed in all cells of these organs and thus is not exclusively localised to cells expressing WT1, suggesting that NAL may play a housekeeping role.

Immunofluorescence indicated that NAL is a nuclear protein and its distribution within the nucleus was interesting as it resembled that of WT1: a speckled pattern within an overall

diffuse staining. Immunocytochemistry confirmed that this protein, like WT1, does indeed co-localise with splicing factors. Thus it will be of interest to see if this protein really is involved in splicing. Unfortunately no further functional assays could be carried out and so future work needs to be carried out to determine its true role and the potential significance of its interaction with WT1.

Chapter 1

Introduction

Chapter 1: Introduction

Wilms' tumour (WT), or nephroblastoma, is a paediatric tumour of the kidney which was first described by Max Wilms in 1899. It is one of the most common childhood cancers occurring in 1 in 10,000 children (Matsunaga, 1981; Pritchard Jones and Hastie, 1990). WT is a striking example of a tumour which has arisen due to development gone wrong and thus serves as a model system for studying the link between development and tumourigenesis.

1.1 Clinical Presentation

The initial clinical manifestation of WT is frequently as a large abdominal mass detected by parents but it can manifest itself with several other symptoms. Most cases of WT are unilateral, where the tumour affects only one kidney, but 5-10% of children can be affected bilaterally. WT is usually diagnosed by the age of six, with the median age of presentation being 3.5 years, although older children and even adults have been known to be affected. Females have a higher incidence especially for the bilateral cases and they are also usually older by an average of six months when they are diagnosed (Breslow and Beckwith, 1988). WT occurs to the same extent world-wide although it does seem to vary between different ethnic groups even within the same geographical area, suggesting that there is a genetic component leading to disease predisposition (Kramer et al., 1984).

WT often comes associated with other congenital diseases, especially in children with bilateral tumours. The first association seen was that of WT arising together with sporadic Aniridia (absence of the iris), Genitourinary abnormalities and mental Retardation, giving rise to the WAGR syndrome (Miller et al., 1964). Beckwith-Wiedemann Syndrome (BWS) is a disease which leads to gigantism, hemihypertrophy, large tongue, hyperplasia of the kidneys, gonadal abnormalities and a predisposition to several neoplasms including WT (Beckwith, 1969). Denys-Drash Syndrome (DDS) is associated with male pseudohermaphroditism, progressive renal insufficiency and Wilms' Tumour (Denys et al., 1967; Drash et al., 1970). The Perlman syndrome involves foetal gigantism, bilateral nephromegaly and cryptorchidism (undescended testis) as well as WT (Perlman et al., 1973).

The treatment of WT has advanced dramatically compared to earlier days when the disease was almost always fatal - there is now a 90% cure rate (Weiner et al., 1998). This enhanced success rate is due to improved clinical management arising from large multicentre trials which were first carried out in the USA by the National

Wilms' Tumour Study (NWTS) (D'Angio et al., 1989). The NWTS classified the extent of the disease into various stages according to the histopathology. The multimodal therapy aims to decrease treatment intensity for the earlier stages, using only surgery and chemotherapy, whilst the more aggressive tumours in later stages will also be given radiotherapy.

1.2 Histology

The malignant transformation occurring in WT is thought to arise from cells within the metanephric blastema which have lost developmental control. These mesenchymal cells can undergo a switch to become epithelial cells and would normally form the nephron after having been induced by the ureteric bud (see section 1.7.1). However, in WT malignant mesenchymal cells have continued to divide beyond their normal point in kidney development, and the progeny cells have attempted to differentiate but cannot do so correctly. Due to this, WT is often classified as triphasic as the tumour can contain blastemal, epithelial and stromal components. WTs can vary according to which component predominates and some even show further multilineage differentiation with the presence of skeletal muscle, bone and cartilage (Hastie, 1994; Miyagawa et al., 1998).

It has been noted that islands of cells resembling metanephric blastema can persist into postnatal life (Beckwith et al., 1990). These lesions, called nephrogenic rests, have been found in only 1% of infant postmortems whereas the incidence in WT kidneys increases to 40% in unilateral and 100% in bilateral cases, suggesting that they are premalignant and predispose to tumour development. The location of these lesions varies among WT patients. One type of rest is called the perilobar nephrogenic rest (PLNR) and this is found near the periphery of the renal lobe, whereas the other type is called intralobar nephrogenic rest (ILNR) and is found within the renal lobe. The ILNR are usually associated with early events occurring in the developing kidney and are often found in children who are diagnosed with WT at an early age, compared to the PLNR which is found in WT diagnosed at a later stage. PLNRs are also found generally in patients that have BWS, whereas ILNRs are normally found in WT from WAGR syndrome (Beckwith, 1997). This observation may suggest that there are different predisposing genes to WT which may be involved in distinct developmental pathways within the kidney and their inactivation may disrupt kidney development at discrete time points.

1.3 Genetic Loci Associated With WT.

1.3.1 The Knudson Model And *WT1*

WT are usually unilateral and sporadic but 5-10% are bilateral and a further 1-2% familial (Breslow and Beckwith, 1988). WT is therefore likely to follow the Knudson model of retinoblastoma, RB (Knudson, 1971), since the latter can also be either unilateral or bilateral, familial or sporadic. Knudson predicted that for RB to develop both alleles of a tumour suppressor gene would have to be mutated. Bilateral cases would carry a germline mutation in one allele of the appropriate tumour suppressor gene, therefore since all cells had a mutation, only one additional genetic hit in the other allele would be required to allow for tumourigenesis. This would explain the incidence of bilateral tumours and the early age of onset. Unilateral cases have a higher incidence but a later age of onset as the child is not predisposed to the tumour and requires two independent genetic mutations. A year later Knudson and Strong proposed a similar "two-hit" model for WT which suggested that there was a specific WT locus and therefore maybe WT, like RB, was caused by loss of function of a tumour suppressor gene (Knudson and Strong, 1972).

Unfortunately the genetics of WT has turned out to be more complicated than that of RB and the proposed model cannot account for all aspects of WT. Unlike the situation with RB, which only has one predisposing gene, WT seems to have multiple genes leading to the disease. One of the WT genes, *WT1*, was localised by cytogenetic studies, which identified chromosome 11 alterations in patients with the cancer. Molecular analysis of deletions associated with WAGR, as well as a homozygous deletion in a WT, mapped the gene to 11p13 (see section 1.4). *WT1* was shown to act as a classical tumour suppressor gene as introduction of various *WT1* isoforms into a WT cell line resulted in the suppression of both *in vitro* and *in vivo* growth (Haber et al., 1993). *WT1* was therefore the gene involved in WT which would follow the Knudson hypothesis. Loss of heterozygosity (LOH) studies, using polymorphic markers to the short arm of chromosome 11, had also been used to determine whether this chromosomal region was regularly disrupted in Wilms' tumours. In chromosomal regions showing LOH, one allele of a particular locus is lost and this can lead to tumour formation if the locus affected carries a tumour suppressor gene and the other allele had already been mutated. LOH can arise due to mitotic recombination; chromosome loss followed by re-duplication of the sister chromosome; or deletion of a region of one of the two sister chromosomes. These LOH studies only detected such allele losses in 30-40% of all WT cases studied, and out of these, a few were restricted to the 11p13 region, but most involved the 11p15

region (Henry et al., 1989; Koufos et al., 1989; Reeve et al., 1989). These two chromosomal regions on 11p must therefore carry two genes involved in WT.

1.3.2 The *WT2* Locus

The *WT2* gene at locus 11p15 has still not been cloned although there are candidate genes. As the BWS and the *WT2* gene are present on the same chromosomal region, it is not yet known whether they are the same gene or two separate genes which are closely linked. Interestingly the LOH in the tumours all seem to be from the maternal chromosome (Schroeder et al., 1987), raising the possibility that perhaps the gene could be imprinted. Also of interest is the fact that some sporadic BWS patients have chromosome anomalies where there is uniparental paternal disomy (UPD) for 11p15, and it is these patients that frequently go on to get embryonal tumours (Henry et al., 1991). The insulin-like growth factor II gene, *IGF2*, has already been mapped to 11p15 and is therefore a good candidate, especially as it is known to be imprinted with only the paternal copy of the gene active (DeChiara et al., 1991; Ohlsson et al., 1993). *IGF2* is a growth factor highly expressed during the embryonic development of many organs, including the kidney. *IGF2* has been shown to have increased expression in WTs (Reeve et al., 1985) and this could be because in cases where there is UPD there would now be two active copies of the paternal gene thus leading to *IGF2* overexpression and hence overgrowth symptoms. There have also been reports where there is biallelic expression of *IGF2* in WTs, occurring due to loss of imprinting (LOI) (Ogawa et al., 1993). This again would lead to overgrowth and tumours. There is, however, no direct evidence that *IGF2* is the second WT gene, and there have been no specific mutations of the *IGF2* gene reported in WTs.

There is another imprinted candidate gene at 11p15, *H19*. This gene is maternally active and can be transcribed to form RNA but it is not translated (Bartolomei et al., 1991). It has been shown that overexpression of this RNA can suppress the tumorigenicity of a rhabdomyosarcoma cell line (Harrington et al., 1993) and thus may be a tumour suppressor gene. Therefore in UPD or in loss of the maternal copy of 11p15 in tumours would mean that no *H19* is transcribed which would allow for tumour formation. More recent findings, however, suggest that *H19* is not a tumour suppressor gene (Lustig-Yariv et al., 1997) and thus the role of *H19* is not really known. There is no evidence to show that *H19* is the *WT2* gene.

Another imprinted gene associated with BWS, is the *p57kip2* gene. This gene maps to 11p15.5 and is a cyclin-dependent kinase inhibitor that arrests cells in G1. If the

maternal allele is lost there would be a decrease in p57kip2 expression, due to the gene being paternally imprinted, leading to an overgrowth phenotype as there would no longer be negative regulation of the cell cycle. However, in WT there seems to be no altered expression of p57kip2 compared to normal tissue, suggesting that this gene is not the *WT2* gene (Overall et al., 1996).

The efforts to clone the *WT2* gene have been hampered because of the large size of the locus, 10MB. However recently Karnik *et al.* have narrowed this locus down by using high density marker LOH analysis of 11p15.4-15.5, and have refined the *WT2* locus to 800kb within 11p15.5, but they also think there is a second distinct locus in that area predisposing to WT (Karnik et al., 1998). The genetics at the 11p15 locus remain unclear and therefore the search for these genes continues.

1.3.3 Additional WT Genes

Significantly LOH at chromosome 11 only accounts for 30-40% of WT suggesting that there must be more genes predisposing to the disease. There have been reports on LOH for markers at 16q in approximately 20% of WT cases (Maw et al., 1992), and a further 12% of cases with LOH at 1p. However there are other tumours which also have LOH at these loci and therefore these loci may be associated with tumour progression (Grundy et al., 1994).

Familial WT is very rare, with only 1% of cases having hereditary transmission. It is inherited as an autosomal dominant trait with incomplete penetrance (Breslow and Beckwith, 1988; Breslow et al., 1996). Out of these only 3 cases have been linked to chromosome 11p13 or 11p15, and further genetic linkage analysis excluded the whole of chromosome 11 from transmission of disease susceptibility (Grundy et al., 1988; Huff et al., 1988), suggesting that there must be at least a third WT locus. The fact that familial WT is so rare complicates the search for a familial WT gene.

However Rahman *et al.* described a family in Canada which showed linkage of WT to chromosome 17q12-21 (Rahman et al., 1996). Surprisingly there is no LOH at this locus and therefore *FWT1* cannot be acting as a tumour suppressor gene in the classical sense. Perhaps a mutation in *FWT1* may be sufficient to start some familial WTs but then the maintenance of the neoplasm becomes independent of the *FWT1* (Rahman et al., 1997). Further work has shown that in other pedigrees, there is exclusion of linkage to 17q12-q21 (Huff et al., 1997), and therefore maybe there is more than one gene which predisposes to WT in large pedigrees. Subsequently McDonald *et al.* found linkage of familial WT to chromosome 19q in five families studied, and they have called this the *FWT2* locus. However they show that there is

another site on 19q different to *FWT2* which may be acting in two different families (McDonald et al., 1998). This work taken together shows that the etiology of WT is genetically heterogenous.

1.4 Cloning The Wilms' Tumour Gene, *WT1*

As discussed in section 1.3, cytogenetic studies on patients who had the WAGR syndrome were crucial in starting to pinpoint the *WT1* locus, as it was noted that these patients had deletions at 11p13 (Francke et al., 1979; Riccardi et al., 1978). Molecular analysis of a large number of WAGR-associated deletions allowed the *WT1* gene to be mapped to a region within 11p13 (Compton et al., 1988; Rose et al., 1990). It was then shown that the *WT1* locus was separate to that causing aniridia as there were cases of sporadic aniridia which did not have WT, and these patients were not associated with the *WT1* locus but instead the one lying next to it, *PAX6*, (Davis et al., 1988). A homozygous deletion in a sporadic WT eventually narrowed the *WT1* region to 350 kb (Lewis et al., 1988). The *WT1* gene, within this deleted region of 11p13, was positionally cloned independently by two different groups. Call *et al.* used a library of human genomic DNA derived from a somatic cell hybrid to identify clones that were homozygously deleted in sporadic WTs. The genomic clones were then used to isolate cDNAs encoding the *WT1* gene (Call et al., 1990). The gene was also isolated by Gessler *et al.* but they used a different technique called chromosome jumping based on the presence of CpG islands which occur at 5' ends of transcription units (Gessler et al., 1990).

1.5 Structure Of *WT1*

The *WT1* gene spans 50kb of DNA and has 10 coding exons (Call et al., 1990; Gessler et al., 1992; Haber et al., 1991). Northern blot analysis using a *WT1* cDNA showed that the transcript formed is 3kb (Call et al., 1990; Pritchard Jones et al., 1990) which can encode up to 16 different isoforms (Schedl and Hastie, 1998) by means of an alternative start site (Bruening and Pelletier, 1996), alternative splicing (Haber et al., 1991; Hsu et al., 1992) and RNA editing (Sharma et al., 1994), figure 1.1. These different isoforms therefore produce proteins that range from 50-55kD (Bruening and Pelletier, 1996; Morris et al., 1991) depending on which combination of start and splice sites are present. The same reading frame is maintained for each of the different isoforms. There are two alternative splice sites: the first includes/excludes 17 amino acids (17aa) encoding exon five; and the second adds/removes the three amino acids lysine, threonine and serine (KTS) between exons 9 and 10. The resulting isoforms will be named according to which

combination of splice sites are selected such that if exon 5 and the KTS do not get spliced out, the protein will have the +17aa and the +KTS and will be called +/+. If exon 5 and the KTS do get spliced out, the protein will be -17aa and -KTS therefore named -/-. If the protein has +17aa but -KTS it will be called +/- and if it has the 17aa spliced out but contains +KTS it will be called -/+. The ratio of each of the four transcripts in the developing mouse kidney is 8.3 +/+; 3.8 +/-; 2.5 -/+; 1 -/- (Brenner et al., 1992), and the ratio seems to remain constant throughout development (Haber et al., 1991). It is not yet known whether these isoform ratios also hold true at the protein level.

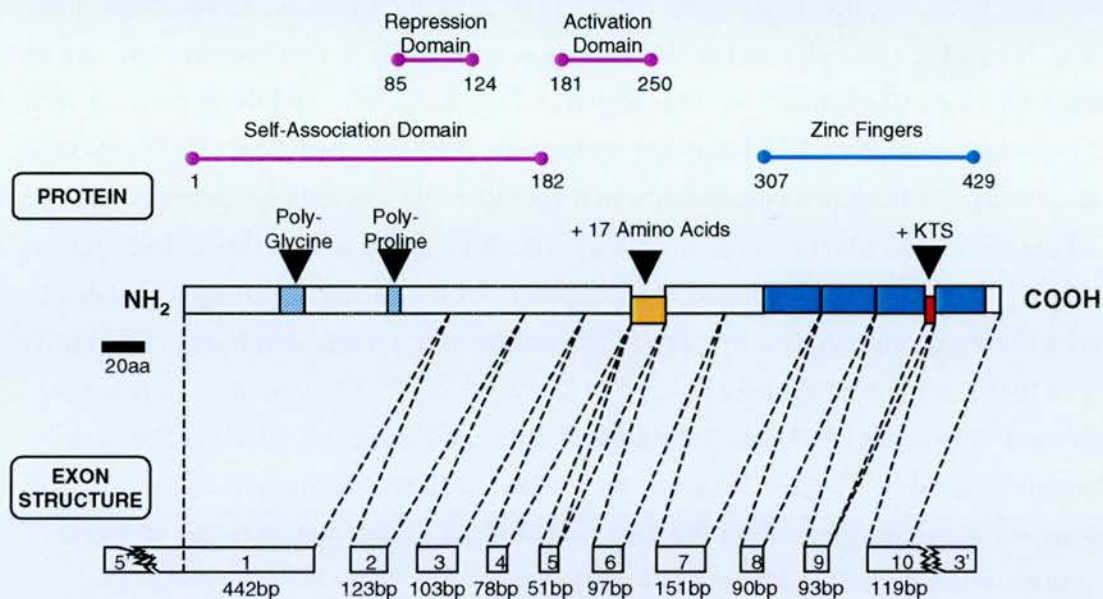


Figure 1.1: Schematic of the WT1 gene and protein structure. Adapted from Huff and Saunders, 1993 and Menke et al 1998.

By studying the conservation of the WT1 cDNA throughout evolution, the importance of the different isoforms and domains seen in the protein can be determined. Kent *et al.* isolated full length cDNAs from chickens and alligators and partial clones from fish and amphibia. The zinc finger region between all vertebrates studied, including human, mouse and rat are 99% identical, even to the extent of the +/- KTS site being conserved. Recently Miles *et al.* have shown that in the pufferfish, *Fugu rubripes*, the 3 amino acids are KPS not KTS, but they can still be spliced out (Miles et al., 1998). This suggests that the alternative splice site and the zinc fingers have a very important role for the function of WT1 in all these animals. The rest of the protein has 87% homology but interestingly the 17aa alternative splice site, the proline rich domain and the RNA editing are not conserved between

all vertebrates, only among mammals (Kent et al., 1995; Miles et al., 1998). The other vertebrates therefore do not have as many isoforms of WT1 as mammals and this could reflect the fact that mammals have a different, more complex genitourinary system.

The predicted protein sequence shows that WT1 has several motifs which could help in the functional characterisation of the protein. The most obvious motif is that of 4 zinc fingers (ZF) of the Krüppel type at the C-terminus (Call et al., 1990). The last three ZFs show close homology to those found in the EGR1 family of transcription factors and therefore possibly form a DNA binding domain. The N-terminus has a proline-glutamine rich domain which has also been found to be common to transcription factors (Mitchell and Tjian, 1989). Another motif which has been found in the WT1 protein is a potential leucine zipper (Madden et al., 1993) and this motif is frequently associated with protein-protein interactions. Finally using structural modelling Kennedy *et al.* have recognised a possible RNA recognition motif (RRM) suggesting that WT1 may bind RNA (Kennedy et al., 1996). From these motifs it seems that WT1 possibly acts as a transcription factor, but since there is a likely RRM domain there could be other potential roles.

1.6 Functions Of WT1

1.6.1 WT1 As A Transcription Factor

DNA-binding specificity of WT1

The structure of the WT1 protein has helped tremendously in trying to ascertain its function. The zinc fingers of WT1 are of the type where a zinc ion is co-ordinated by two cysteine and two histidine residues (C2H2 class). This motif has been found in site-specific DNA binding proteins, and the most highly homologous protein to WT1 in the zinc finger region is EGR1. Structural analysis of EGR1 showed that the protein fits into the major grooves of DNA and each of the fingers interact with specific nucleotides within the 9 base pair sequence GCG GGG GCG (Pavletich and Pabo, 1991). As the last three zinc fingers of WT1 are so homologous to the EGR1 protein, Rauscher *et al.* investigated whether the same consensus sequence could be bound by WT1. Using bacterially expressed WT1 -KTS zinc fingers and gel retardation assays they showed that this zinc finger region of WT1 could indeed bind to this sequence but they also showed that WT1 could bind to slightly different sequences (Rauscher et al., 1990). As WT1 can bind to the same consensus sequence as EGR1, it may seem likely that these two proteins compete for the same binding site. However, Hamilton *et al.* have shown that the presence of the first zinc

finger in WT1, which is not so closely homologous to the EGR1 zinc fingers, apparently affects the DNA binding properties of WT1, generating a GC-rich 12bp consensus recognition site (Hamilton et al., 1995).

The presence of the three amino acids, KTS, between zinc fingers three and four causes the spacing between these two fingers to alter such that the EGR1 consensus sequence can no longer be bound with such high affinity. There is therefore a difference in DNA binding specificity between the + and - KTS isoforms. The +KTS fingers can however still bind to DNA but the sequences they bind to differ to those bound by -KTS fingers (Bickmore et al., 1992). Hence it seems likely that by producing alternative isoforms, WT1 can bind to different target sequences and therefore potentially to different target genes. However, it remains unclear as to what the real consensus sequences are for the different WT1 isoforms.

Potential target genes for WT1

The structure of WT1 and the fact that it can bind to DNA suggests that WT1 acts as a transcription factor. With this in mind it was then shown that WT1 expressed from a cytomegalovirus (CMV) expression vector could repress transcription from a synthetic chloramphenicol acetyltransferase (CAT) reporter construct containing three Egr1 sites (Madden et al., 1991). The repressor domain of WT1 was mapped to amino acids 85-124 and Madden *et al.* showed that if they fused this domain to the zinc fingers of EGR1, the heterologous protein would now function as a transcriptional repressor too. Thus unlike EGR1 which activates transcription, WT1 seems to act as a transcriptional repressor. The 17aa of exon five also seem to have a function in increasing the degree of repression, as it was shown that the WT1 +/+ protein suppresses the activity of the WT1 promoter about 25 fold better than the WT1 -/+ protein (Rupprecht et al., 1994).

In an attempt to identify authentic cellular targets, the *IGF2* promoter was also used in these experiments. *IGF2* had been reported to be overexpressed in WT and the BWS, which has paternal duplications of the chromosomal region containing *IGF2*, also has a predisposition for WT. The promoters that drive *IGF2* expression are GC rich and contain several potential WT1 binding sites and for all these reasons, it was thought that perhaps the *IGF2* gene could be a target for WT1 transcriptional repression. In these *in vitro* experiments, WT1 did repress transcription directed from the *IGF2* promoter (Drummond et al., 1992). DNase I footprinting and gel shift analyses showed that WT1 did indeed bind directly to the *IGF2* promoter.

Gene	Reference
WT1	(Hofmann et al., 1993)
EGR1	(Rauscher et al., 1990)
IGF2	(Drummond et al., 1992)
IGF1R	(Werner et al., 1993)
PDGF-A	(Wang et al., 1992)
EGFR	(Englert et al., 1995a)
RAR α	(Goodyer et al., 1995)
PAX2	(Ryan et al., 1995)
Inhibin- α	(Hsu et al., 1995)
Bcl2	(Hewitt et al., 1995)
c-myc	(Hewitt et al., 1995)
ODC	(Moshier et al., 1996)
CSF-1	(Harrington et al., 1993)
TGF β	(Dey et al., 1994)
Nov	(Martinerie et al., 1996)
Midkine	(Adachi et al., 1996)
p21	(Englert et al., 1997)

Table 1.1: Potential target genes regulated by WT1.

Since then, there have been several other potential target genes whose expression can be regulated by WT1 as shown in the table 1.1. For a full review of potential targets and WT1 DNA-binding sequences see Menke et al., 1998(b). Although it seems that there is a large list of potential targets for WT1, care must be taken because all of these assays have been carried out *in vitro* and not under physiological relevant conditions. Also it may be misleading to suggest that WT1 is always acting a repressor because it has been found that under certain conditions WT1 can activate transcription. For example the reporter constructs used seems to affect the activity of WT1, as has been highlighted by the case of the *PDGF-A* gene. If the 5' residues to the start site of the *PDGF-A* gene were used, WT1 acted as a activator but if the 5' and 3' residues of the *PDGF-A* promoter were used, WT1 now seemed to repress transcription (Wang et al., 1993). The activation domain of WT1 seems to map to position 181-250 and it can act independently of the repression domain. The activity of WT1 also seems to depend on the cell line used (Moshier et al., 1996) as well as on the expression vectors used to drive expression of WT1 (Reddy et al., 1995a).

WT1 therefore seems to be able to both activate some target genes and repress others. This could reflect the role WT1 plays in kidney development, as it may repress genes involved in proliferation and activate those that instead cause the cell to differentiate into epithelial cells. WT1 also seems to regulate genes that are involved in apoptosis, and so it appears that loss of WT1 function deregulates multiple pathways that maintain cell growth in check and the loss of this regulation can result in tumour formation.

1.6.2 WT1 As A Post-Transcriptional Regulator

WT1 is associated with splicing factors

The initial experiments carried out all pointed towards WT1 being a classical tumour suppressor gene whose function was to regulate the transcription of genes involved in proliferation. It has since been discovered that perhaps there is a further function for WT1 as a post-transcriptional regulator. The initial studies came from Larsson *et al.* where the subnuclear localisation of WT1 was examined. They looked at the nuclear staining pattern of WT1 in a mesonephric cell line (M15) which expresses WT1 endogenously. They found that WT1 staining not only gave a diffuse pattern, as is found for other transcription factors, but there was also a characteristic speckled pattern (Charlieu *et al.*, 1995; Larsson *et al.*, 1995). These 20-50 distinctly stained irregular foci, which exclude the nucleolus, have been previously documented for the subnuclear localisation of splicing factors (Spector *et al.*, 1991). These speckles are interchromatin granule clusters, and it is believed that they are either storage sites for splicing factors, or sites where the splicing factors are pre-assembled into spliceosomes, a multicomponent complex involved in splicing. Splice factors are then recruited from these speckles to the perichromatin fibrils, which are the sites where active transcription occurs, where they then remove the introns from the nascent pre-mRNA (Misteli *et al.*, 1997; Spector, 1996). By transfecting each isoform in turn into COS cells Larsson *et al.* found that it was mainly the -KTS isoforms that produced the diffuse pattern whereas the +KTS isoforms went predominantly into the speckles. By using double labelling experiments they showed that the speckles where WT1 was present overlapped with the speckles stained by splicing factors, and so the idea that WT1 might play a role in splicing arose.

Englert *et al.* carried out similar experiments and they seem to disagree with the idea that WT1 has a role in splicing as they think that the speckled structures for WT1 are different to the structures that contain pre-mRNA splicing factors. They showed, using an inducible cell line, that some of WT1 +KTS did co-localise with a

spliceosome assembly factor, SC35, but a lot of the WT1, as well as the mutant forms of the protein, seemed to go into other large foci which did not co-localise with SC35 (Englert et al., 1995b). They suggest that these sites are merely storage sites for mutant WT1 protein or for the WT1 protein that has a decreased DNA-binding affinity. Although it is possible that they may act as storage sites, as it is known that both active and inactive splice factors are stored in the speckles (Gama-Carvalho et al., 1997), it seems unlikely that all the +KTS isoforms, which are the most abundant forms of WT1, would be stored due to its decreased DNA-binding affinity and for it not to have any other function. Further evidence to suggest that at least some of the WT1 present in the speckles may have a role in splicing came from showing that WT1 could directly interact with splicing factors as an antibody directed against splicing factors could co-immunoprecipitate WT1 *in vivo* (Larsson et al., 1995).

There is other evidence to suggest that WT1 is indeed involved in splicing. Continuing on from Larsson's work, Davies *et al.* wanted to investigate whether WT1 does directly interact with the splicing machinery and if so with which components. They have now shown that, both in the yeast two hybrid system and *in vivo*, WT1 can bind to U2AF65, a protein required for the binding of the U2 snRNP to the intron branch point adjacent to the 3' splice site (Zamore et al., 1992), and may therefore have a role in 3' splice site selection (Davies et al., 1998). Furthermore they showed that the +KTS isoform bound to this splicing factor with the greatest affinity, supporting the idea that the WT1 +KTS isoform, which is the predominant form found in speckles, functions in splicing. They also provide evidence that WT1 can become incorporated into spliceosomes *in vitro*.

WT1 can bind to RNA

It has been suggested that WT1 can not only contact DNA through its zinc fingers, but it can also bind to RNA. WT1 has been shown to bind to the promoter sequence of the *Igf2* gene and also to sequences within exon 2, which is one of three alternative 5' untranslated leader exons, at the RNA level (Caricasole et al., 1996). They demonstrated that for this RNA binding, it was the zinc finger region which was contacting the transcript and they showed that it was specific to WT1 as the closely related EGR1 protein could not do so. More specifically they established that it was zinc finger 1 that had a more significant role in RNA binding, compared to zinc finger 4. Bardeesy *et al.* also showed that WT1 could bind to RNA. Using an RNA selection method, they pulled down 3 groups of RNA ligands which are specifically recognised by WT1 but they show that zinc fingers 2-4 and not zinc

finger 1, were important for binding (Bardeesy and Pelletier, 1998). Thus it seems that different combinations of fingers can again select for different RNA binding sites. There have been other reports showing that WT1 may have the potential to bind to RNA. Kennedy *et al.* used molecular modelling to show that WT1 has an RNA recognition motif (RRM) mapping at the N-terminus of the protein (Kennedy *et al.*, 1996). This region is totally separate from the zinc fingers, suggesting that WT1 can contact RNA by more than one mechanism. These findings that WT1 can bind to RNA are carried out *in vitro* and it is not known how this reflects the *in vivo* situation. However this is not the first example of a protein binding to both DNA and RNA. TFIIIA, which is another zinc finger protein can also regulate gene expression by both transcriptional and post-transcriptional mechanisms (Ladomery, 1997). Nevertheless further work needs to be carried out to determine what the real RNA targets for WT1 are, and the consequences of such an interaction.

1.6.3 WT1 is involved in cell growth regulation and apoptosis

The importance of the WT1 function in suppression of cell growth was shown when WT1 was identified as a tumour suppressor gene. In these experiments, wild type *WT1* was reintroduced into a Wilms' tumour cell line expressing an aberrantly spliced *WT1* transcript, and shown that it could suppress cell growth (Haber *et al.*, 1993). Kudoh *et al.* also showed that microinjection of a WT1 cDNA into cells that were at early G₁ stage, blocked cell cycle progression into S phase (Kudoh *et al.*, 1995). This cell cycle arrest by WT1 seems to be associated by the induction of p21, independent of p53 (Englert *et al.*, 1997). WT1 and p21 are found in the same developing structures within the kidney and perhaps they work together to stop cells from proliferating and start differentiating (see section 1.7.1).

Some cells undergo apoptosis after their G₁ arrest by WT1, as shown by the induction of WT1 (especially the -KTS isoforms) in U2OS and Saos-2 cells (Englert *et al.*, 1995a). They also showed that the induction of cell death by WT1 was associated with reduced expression of EGFR. The role of WT1 in apoptosis is important as it is part of the developmental programme of the kidney (Koseki *et al.*, 1992; see section 1.7.1).

If, however, WT1 is expressed at lower levels, it represses p53-mediated apoptosis. At these lower levels, WT1 cannot repress EGFR and so does not induce apoptosis, but instead binds to p53 and inhibits p53-mediated apoptosis due to ionising radiation (Maheswaran *et al.*, 1995). This could explain why mesenchymal cells in

the developing kidney which normally undergo an epithelial transition (see section 1.7.1) actually undergo apoptosis in *wt1*-null mice (Kreidberg et al., 1993).

Whether a cell undergoes apoptosis depends on several signals it receives, and the levels of WT1 seem to influence whether apoptosis will occur. Apart from the pathways already mentioned, the regulation of bcl-2 and c-myc by WT1 will also determine whether a cell goes into apoptosis (Hewitt et al., 1995). The role of WT1 in growth arrest and apoptosis is known to be important as it was shown that the adenovirus can transform cells by expressing the E1B 55 K protein which sequesters WT1, along with p53, within a cytoplasmic body thus not allowing them to carry out their function (Maheswaran et al., 1998a).

1.6.4 WT1 Can Interact With Several Cellular Proteins

In order for WT1 to carry out its various roles it must interact with the cellular machinery. Whether WT1 activates or represses transcription; is involved in post-transcriptional modifications; or represses cell growth and induces or inhibits apoptosis will depend to a certain extent on its interaction with other cellular proteins. This was first seen with p53, which was shown to bind to WT1 by immunoprecipitation experiments and by doing so affecting the transcriptional activity of WT1 as shown by transfection assays. In the absence of p53, WT1 activates an EGR1 reporter but in the presence of p53 it acts as a repressor of EGR1 transcription (Maheswaran et al., 1993). The binding of WT1 to p53 also affects apoptosis as described earlier.

It has also been shown that WT1 can bind to more proteins than just p53. WT1 itself can dimerise as shown using the yeast two hybrid system. The domain required for this self-association has been mapped to position 1-182 (Englert et al., 1995b). The human prostrate apoptosis response-4 (par-4) protein also binds to WT1, as shown by *in vitro* and *in vivo* experiments (Johnstone et al., 1996). Par-4 is a novel protein required for apoptosis in prostatic cells. By binding to WT1, par-4 inhibits its transcription activation and augments the repression activity as shown by transient transfection studies. Par-4 does not affect the DNA-binding activity of WT1 so it has been suggested to function by bringing another repressor domain to the promoter. Par-4 also seems to be able to rescue the growth suppression caused by WT1 and sensitises cells to apoptosis (Sells et al., 1997). Thus WT1 is a protein partner to par-4 and attenuates its function to block cell growth and apoptosis. Increased par-4 expression would reverse the WT1 effects.

WT1 can also bind to UBC9, the human ubiquitin conjugating enzyme (Wang et al., 1996). The exact effect this has on the function of WT1 is not clearly understood but it may affect the role WT1 plays in the cell cycle. WT1 (-/+) and (+/+) can block the cell cycle progression into S-phase by downregulating the activities of CDK2 and CDK4. This block can be overcome by overexpression of cyclin E/CDK2 as well as cyclin D1/CDK4 (Kudoh et al., 1995). UBC9 probably acts by the degradation of S- and M-phase cyclins (Seufert et al., 1995) and so has a role in cell cycle progression. Therefore if WT1 binds to UBC9 it may impose a cell cycle block by affecting the fate of the S- and M-phase cyclins that are degraded by ubiquitination.

The inducible heat shock protein 70, Hsp70, is another protein that WT1 binds to, as shown by co-immunoprecipitation experiments (Maheswaran et al., 1998b). Hsp70 binds to the extreme amino terminus of WT1, and it may ensure correct folding of this domain, enhancing the functional properties of WT1. Binding of Hsp70 by WT1 appears to be necessary for the latter to be able to induce p21 expression and mediate a G₁ phase cell cycle arrest.

Ciao1, identified by its ability to bind to WT1, is another protein that can mediate the transcriptional activation of WT1 (Johnstone et al., 1998). Unlike par-4 though, it does not affect the repression activity. Ciao 1 is a novel protein and is a member of the WD40 family. Proteins containing the WD40 repeats have several functions including cell cycle regulation, RNA splicing, and transcription. Thus although it has been shown by transient transfections that Ciao 1 can affect the transcriptional activity of WT1, it may also affect some of the other roles of WT1.

The steroidogenic factor 1 (SF-1) protein has also been shown to bind to WT1, and the binding is isoform specific as SF-1 only binds to the -KTS isoforms. Together they are involved in promoting Müllerian inhibiting substance (MIS) expression which determines the fate of the gonad (Nachtigal et al., 1998; see section 1.7.2) and thus their interaction is vital for gonadal development.

It is clear that there must be other cellular proteins with which WT1 interacts for it to carry out its many roles in transcriptional control, splicing, apoptosis and cell cycle regulation.

1.7 Expression Pattern Of WT1

The expression pattern of WT1 has been looked at in detail to try to gain further insights into its role in development and tumourigenesis. Northern blot analysis

initially showed that *WT1* could be detected in foetal kidney, spleen, testis, ovary and some weak expression in the brain (Pritchard Jones et al., 1990). More detailed analysis examined the spatial and temporal expression pattern of *wt1 in situ* (Armstrong et al., 1993). They showed that *wt1* is first expressed in the intermediate mesoderm but is then more strongly expressed in the urogenital system during organogenesis. *Wt1* is also expressed in the mesothelium, an epithelial tissue which surrounds the internal organs of the body. There is also some expression of *wt1* along the spinal cord and a region of the brain; adult bone marrow (Fraizer, 1995); lymph nodes; and peripheral blood (King-Underwood et al., 1996). Apart from the latter regions, *wt1* is present in tissues of mesodermal origin that undergo a mesenchyme to epithelial transition. This expression pattern seems to fit in with the tissues affected when *WT1* is mutated.

1.7.1 The Expression Of *WT1* In The Developing Kidney And The *Wt1* Knockout Mouse

The expression pattern of *WT1* has been studied extensively in the developing kidney (Armstrong et al., 1993; Pelletier et al., 1991b; Pritchard Jones et al., 1990). Its pattern of expression suggests that *WT1* has a role in at least three stages of renal development: both at the onset and during progression of nephrogenesis and also in the maintenance of mature podocyte function.

The development of the kidney has been studied in both humans and mice and they follow similar pathways. A brief description of mouse kidney development will be described to show the importance of the different areas where *wt1* is expressed. For reviews on kidney development see Davies, 1996; Davies and Bard, 1996; Lechner and Dressler, 1997; Saxen, 1987.

During embryonic day 8 and 10 (E8 and E10) the pronephros and mesonephros are formed, and these are transient precursors for the permanent kidney. These rudimentary kidneys consist of the nephric, or Wolffian, duct and some tubules. By E10.5 the ureteric bud (which is the primordium of the renal collecting duct) evaginates from the most caudal end of the duct and invades the blastema of the metanephric mesenchyme. At this stage, the mesenchyme are the only cells that express *wt1* albeit at a low level (figure 1.2). The metanephric kidney then starts to form through the reciprocal interactions of these two tissues. The interaction induces the formation of a metanephric mass over the growing ureteric bud, which in turn causes the latter to branch within the metanephric blastema. The blastema is induced to condense at the tip of each branch, which in turn causes further branching and thus a wave of nephrogenesis occurs throughout the developing

kidney. The expression of *wt1* is upregulated in the condensing mesenchyme. The condensed mesenchyme then undergoes a burst of proliferation before differentiating into epithelial cells that form the renal vesicle, which continues to express *wt1*. At this stage there is a wave of regression which removes the pro- and mesonephros.

The renal vesicle then undergoes morphological changes and first forms a comma shaped body, so called because cells further from the collecting duct become elongated and a slit forms within the vesicle. Another slit is then formed but this time at the distal end, near the collecting tubules and this structure is now called an S-shaped body. At this point *wt1* expression is confined to the proximal end of the vesicle. The distal end of the S-shaped body then joins the collecting duct and a series of collecting tubules are formed. The proximal end of the vesicle becomes the Bowman's capsule, the inner cells of which form the glomerular podocytes and this is where *wt1* expression is confined to. *Wt1* expression in the glomerulus will remain postnatally but at a lower level. This is the only place in the mature kidney where *wt1* is expressed as any remaining mesenchyme which has not been induced by the ureteric bud undergoes apoptosis (Koseki et al., 1992).

As can be seen from these expression studies, *WT1* must have an important function in the formation of the nephron. Mutations in *WT1* do lead to developmental abnormalities within the kidney, but there are no known humans with a null mutation for *WT1* which would prove that *WT1* is essential for early stages of nephrogenesis. Kreidberg *et al.* therefore used gene targeting to produce *wt1* null mice (Kreidberg et al., 1993). These mice, which die at mid-gestation due to heart defects, have no metanephric kidneys at all as the metanephric mesenchyme fails to differentiate and therefore undergoes apoptosis. The ureteric bud does not grow out and degenerates. This therefore proves that *wt1* must be essential for kidney development and suggests it may have a role in producing the signal which is thought to be secreted to induce the ureteric bud. The exact nature of this signal is not yet known and in fact there are many gene products produced during kidney development but it is not known how they all function and interact with each other to form the kidney (Davies and Brandli, 1994). One potential candidate signalling molecule required for ureteric bud formation and branching is the glial cell line derived neurotrophic factor (GDNF). GDNF is expressed by the mesenchyme and it binds, via the receptor tyrosine kinase Ret, to the tip of the ureteric bud, causing branching during metanephric development (Pichel et al., 1996).

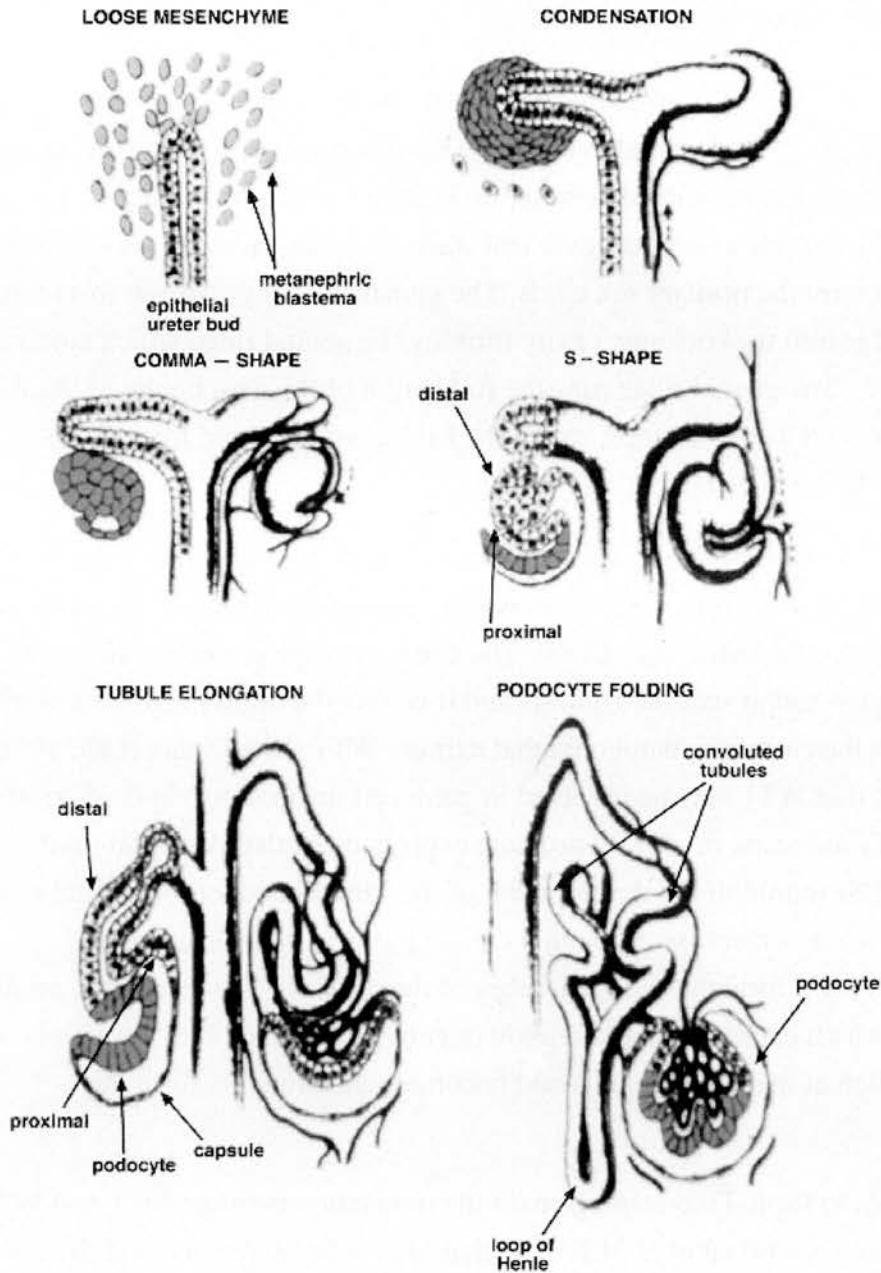


Figure 1.2: Diagram illustrating the expression pattern of WT1 during nephrogenesis. WT1 is expressed at low levels in the mesenchyme (light shading), and the expression is then upregulated during the condensing of the mesenchyme around the ureteric bud (stronger shading). WT1 expression then remains on in the comma-shaped bodies, as well as the S-shaped bodies except that it becomes localised to the proximal end. The expression of WT1 finally becomes confined to the podocyte layer of the glomerulus. (Figure taken from Hastie, 1994)

1.7.2 *WT1* Expression in the Gonads

The *wt1* knockout mouse showed that WT1 was also essential for gonad development as the gonads failed to develop (Kreidberg et al., 1993). This is not surprising as gonad development is linked to kidney development. Initially gonadal development is the same for both sexes and starts as a thickening of the coelomic epithelium to form the primary sex cords. The gonad rapidly grows out from the urogenital ridge into the coelomic cavity forming the genital ridge which continues to express *wt1*. The genital ridge runs the full length of the mesonephros. As this is occurring a second duct is formed, called the Müllerian duct, and like the mesonephric duct it does not express *wt1*.

At around E13 the differentiation programme for the female and male gonads differ. In the testis the seminiferous cords form within the mesenchyme and they consist of epithelial Sertoli cells (McLaren, 1998). The cords themselves enclose the germ cells, which give rise to spermatogenesis and it is only the Sertoli cells surrounding the cysts with the early spermatogonia that express WT1 (Rio-Tsonis et al., 1996). This suggests that WT1 may be involved in germ cell maturation. In the Sertoli cells SF-1 and WT1 associate *in vitro* to promote expression of the Müllerian inhibiting substance (MIS) to inhibit the development of the Müllerian duct (Nachtigal et al., 1998) which would otherwise, through a default pathway (Ramkissoon and Goodfellow, 1996), form the fallopian tubes of the ovaries. The germ cells produce testosterone which causes the mesonephric duct to differentiate into the Vas Deferens, which at the level of the gonad becomes convoluted to form the epididymis.

For the ovaries to form, Dax-1 antagonises the interaction between SF-1 and WT1 in the Sertoli cells (Nachtigal et al., 1998) so that MIS is not expressed and the default pathway can proceed. In the ovary there are no seminiferous tubes and germ cells become surrounded by granulosa cells which express *wt1*. The germ cells of the ovary produce oestrogen which causes regression of the entire mesonephric duct. It also causes the Müllerian duct to differentiate into the oviduct. The caudal ends of the oviducts then fuse to form the uterus.

WT1 therefore plays an important role in gonad development and this can be seen both in humans, where if *WT1* is mutated gonadal dysgenesis such as that found in DDS occurs, and in the knockout mouse where no gonads develop. However because after gonad development has occurred *WT1* is still expressed both in Sertoli

cells and in granulosa cells, it suggests it may also play a role in both sperm and follicle development (Hsu et al., 1995; Rio-Tsonis et al., 1996).

1.7.3 Regulation of *WT1* Expression

Unlike many other ubiquitously expressed tumour suppressor genes, *WT1* has a specific expression pattern which is temporally and spatially restricted (Armstrong et al., 1993). The expression of *WT1* has to be tightly regulated as its misexpression could lead to disastrous consequences. For instance *WT1* can increase the tumour growth rate of adenovirus-transformed baby rat kidney cell line (Menke et al., 1996) and if *WT1* is expressed from a CMV promoter in transgenic mice such that it would be ubiquitously expressed, it causes embryonic lethality (Menke et al., 1998a).

The upstream regulatory elements of the *WT1* gene control region have been characterised and the promoter seems to be TATA-less and GC-rich (Hofmann et al., 1993), and it has multiple transcription start sites present (Pelletier et al., 1991b). A far upstream cis-element, approximately 15kb upstream of the transcription start site, has been identified. This element is necessary but not sufficient for *wt1* expression in human foetal kidney cells in culture (Scholz et al., 1997). Hofmann *et al.* (1993) have also shown by DNase I footprinting that there are several sites within the promoter where Sp1 binds and by transient transfections they showed that the *WT1* promoter was activated in response to Sp1. Initially it seemed strange that Sp1 could regulate the expression of *WT1* as Sp1 is ubiquitously expressed and therefore would not be able to produce the specific expression pattern seen for *WT1*. However a more detailed study of Sp1 showed that the levels of Sp1 expression are not constant but that it is expressed most highly in the areas where there is *WT1* expression (Cohen et al., 1997). Hofmann *et al.* also found *WT1* binding sites within the *WT1* promoter and it seems that *WT1* itself can regulate its own expression in a negative feedback loop.

Other genes which seem to affect *WT1* expression are *Pax2* and *Pax8* both belonging to the paired box family of genes and involved in kidney development. The *WT1* promoter has two *Pax2* binding sites and transient transfections show that *WT1* expression is upregulated by *Pax2* (Dehbi et al., 1996). *Pax2* expression is induced in the condensing mesenchyme and it is after this, in the formation of the renal vesicle that *WT1* expression increases. Interestingly it is then that the *Pax2* expression falls. As mentioned earlier, *WT1* can repress the expression of *Pax2* (Ryan et al., 1995), and therefore when *Pax2* expression peaks, it activates *WT1* expression, which in turn turns off *Pax2* expression via a feedback loop. *Pax8*

which is expressed at its highest level during renal vesicle formation has been shown to activate the *WT1* promoter (Dehbi and Pelletier, 1996) and it is at this stage that *WT1* expression peaks too. When levels of Pax8 decrease so do the levels of WT1, which may also be due to WT1 inhibiting expression at its own promoter (Rupprecht et al., 1994).

It has been noted that *WT1* expression can be specifically regulated in haematopoietic cell lines as the promoter has a cell-line specific enhancer within the 3' end of the gene (Fraizer et al., 1994) which is transactivated by GATA-1 (Wu et al., 1995). This is of interest because WT1 is thought to be expressed in organs involved with haematopoiesis.

Although we are beginning to understand a little more about how the expression of *WT1* is regulated it is still not understood how the precise temporal and spatial expression is achieved. It is also not known whether the factors affecting *WT1* expression work together or whether they compete for sites on the promoter. Following isolation of the genomic *WT1* from pufferfish *Fugu rubripes* (Miles et al., 1998) work is in progress to examine the promoter region. As the genome of the fish is a lot more compact, it should highlight those regions which are most important for proper expression. Transgenic mice containing this fish WT1 can be used to investigate whether the fish genomic WT1 clone contains all the regulatory elements in the promoter region for it to be expressed in the appropriate areas in the mammalian kidney, an organ which is very primitive in the lower vertebrates.

1.8 WT1 Involved In Disease

Mutations in *WT1* have been shown to be involved in the development of WT, and it was thought that WT1 was acting as a classical tumour suppressor gene, and patients with the WAGR syndrome would have both copies of the *WT1* gene inactivated. However, extensive studies have shown that fewer than 15% of WT cases actually have *WT1* mutations and that therefore there are other genes, such as the *WT2* and the *FWT1* and *FWT2* (see section 1.3), which predispose to this disease to a greater extent than *WT1* (Gessler et al., 1994). The studies carried out however cannot guarantee that all mutations in WT1 are detected and therefore the precise percentage of WT in which *WT1* is involved in is not known.

There are also cases where only one *WT1* allele is mutated and thus *WT1* would not be acting as a classical tumour suppressor gene. Heterozygosity could, however, still predispose to WT as there may be haploinsufficiency leading to tumourigenesis. However mice lacking one allele of *wtl* do not develop WT (Kreidberg et al., 1993)

which suggests that haploinsufficiency of *wt1* does not lead to WT. On the other hand, there are cases where *WT1* is known to be heterozygously mutated, such as in some cases of DDS. A large proportion of DDS patients have a missense mutations within exon 9 resulting in a replacement of an arginine for a tryptophan at position 394 (Pelletier et al., 1991a). Most other patients have mutations somewhere in the zinc finger region and all of these cause the WT1 protein to lose its ability to bind to DNA (Little et al., 1995). This would suggest that the mutant protein must actively be involved in bringing about the disease because it is known from the mice studies that reduced levels of WT1 alone do not lead to WT. It has been shown that the mutant WT1 protein, which cannot bind to DNA, dimerises to the wild type WT1 protein and inhibits its transcriptional activation (Moffett et al., 1995; Reddy et al., 1995b). The mutant protein therefore acts in a dominant negative fashion.

Frasier syndrome (FS) is a disease which shows the importance of maintaining the correct isoform ratio of WT1 for normal development. FS is a rare, dominant heterozygous disease defined by male pseudohermaphroditism and progressive glomerulopathy but no WT. This disease is thought to arise through a mutation in the intron of exon 9 which would lead to the loss of +KTS and an imbalance of the WT1 ratios, thus showing the importance of all the WT1 isoforms (Barbaux et al., 1997). The phenotype of this disease suggests that the -KTS isoforms, and not the +KTS isoforms, are involved in acting as tumour suppressor genes, as patients with FS do not develop WT.

Apart from these syndromes, mutations in the *WT1* gene are also thought to be involved in other tumour types. The types of tumours that have *WT1* mutations correlate with the expression pattern seen for WT1. For example, WT1 is expressed in the mesothelium and *WT1* mutations have been found in non-asbestos related mesothelioma (Park et al., 1993). However, it should be noted that only 1 mesothelioma out of 100 had WT1 mutations (Langerak et al., 1995). Mutations in WT1 are also thought to be involved in leukemias as WT1 is highly expressed in a range of acute leukemias and haematopoietic cell lines and mutations in the gene have been found to be present in this disease (King-Underwood et al., 1996; Miwa et al., 1992). Ovarian tumours have also been found to contain *WT1* mutations (Bruening et al., 1993). Thus a picture is emerging where WT1 is not just involved in WT and perhaps it plays a bigger role in other tumour types (Little and Wells, 1997).

1.9 Summary and aim of project

Wilms' tumour is a good model to study the link between cancer and development. There are several genes predisposing to the disease, but only the *WT1* gene has been cloned and shown to be directly involved in urogenital development. The gene encodes 16 different isoforms of a protein that has a structure resembling other transcription factors as it has four zinc fingers, three of which are highly homologous to the EGR1 family. The protein also has a proline-glutamine rich domain which again is found in other transcription factors. Indeed, transient transfections have shown that WT1 can either activate or repress certain target genes. It has been noted that WT1 can have different effects in different cell lines and this must depend on the presence of other cellular components which WT1 can interact with.

The structure of the protein, however, also points to a role in post-transcriptional regulation as the zinc fingers are also capable of binding to RNA and there is a possible RNA recognition motif at the N-terminus. The +KTS isoform of WT1 has been found in speckled structures of the nucleus, which are often associated with splicing factors, and WT1 has been shown to bind to splicing factors *in vivo*. Thus the various isoforms seem to play different roles in the cell, although it is not known how much each isoform contributes to the various functions.

Detailed expression studies have shown that *WT1* is expressed mainly in cells that are undergoing a mesenchyme to epithelial transition, although there are other pockets of expression. The main areas where WT1 is found are in the developing kidney and gonads and it shows a highly regulated spatial and temporal expression pattern. Although the *wt1* knockout mice show that WT1 is essential for urogenital development, the expression of *WT1* continues on in some adult structures, such as the Sertoli cells and the follicular cells of the male and female gonads respectively as well as in the podocytes, suggesting that WT1 may play a further role in these structures.

It is thought that WT1 may function by controlling the expression of genes involved in cell proliferation and apoptosis. From the expression pattern seen in the kidney, it is likely that WT1 represses these genes and activates other genes involved in differentiation, thus allowing for the mesenchyme to epithelial transition. *WT1* mutations may lead to tumour formation due to mutant WT1 having lost the ability to bind to DNA and thus it would not be able to repress those genes involved in cellular proliferation leading to mass cellular production without any differentiation.

Although a lot of progress has been made in trying to understand the role of WT1 there is still a lot to learn. It is not yet clear how WT1 brings about its functions and why there are so many isoforms. How does WT1 bring about both activation or repression of certain genes? Which genes are targeted by WT1? Are all isoforms involved in transcription? The environment WT1 is in seems to affect its function. However, it is not yet known how WT1 interacts with other cellular components to bring about the required activity. The potential role of WT1 in splicing is still very controversial. It is not really known if WT1 causes post-transcriptional regulation - does it bind to certain transcripts, and if so which? What happens to any of the transcripts which are bound by WT1? Does WT1 affect the splice site selection? These are all questions that still need to be addressed. It is very important to find out with which other components of the cellular machinery WT1 is interacting. By identifying other proteins which WT1 binds to, we may start to shed some light as to how WT1 acts at the molecular level. This project will therefore use the yeast two hybrid system as a tool to identify WT1-interacting proteins. By knowing the function of the proteins WT1 binds to we may gain some insight into how WT1 can apparently have so many different activities. The different isoforms can be looked at to see if they all bind similar proteins. For example the +KTS isoforms might tend to bind to proteins involved in splicing whereas the -KTS isoforms may bind to proteins which could affect their transcriptional activity. Proteins involved in cell cycle regulation and apoptosis may also emerge as indeed may new proteins which may suggest that WT1 also has further roles other than those already documented. It is only once we know the function of WT1 at the molecular level that we will be able to dissect its true role in urogenital development and its link with tumourigenesis.

Chapter 2

Materials and Methods

Chapter 2: Materials and Methods

2.1 Bacterial Work

2.1.1 Bacterial Strains Used

Bacterial Strain	Genotype
TG1	supE, hsd Δ 5, thi Δ (lac-proAB), F'[traD36, proAB+, lacI ^q Δ M15]
Y1090	supF, hsd R, araD 139 Δ lon, Δ lacU169rpsL, trpC22::Tn10(tetr)pMC9
XL-1 Blue	supE44, hsdR17, recA1 end A1, gyrA46, thi rel A1lac ⁻ F'[proAB+, lacI ^q , lacZ Δ M15, Tn10(tet ^r)]

2.1.2 Media and Additives

All bacterial media was provided by in house stores, unless otherwise stated.

L-Broth - 10g tryptone, 5g yeast extract, 10g NaCl, and 2.46g MgSO₄ in 1 litre H₂O.

L-Agar - 10g tryptone, 5g yeast extract, 10g NaCl, 2.46g MgSO₄ and 15g agar (Oxoid Ltd) in 1 litre H₂O.

CY-TOP -10g casamino acid, 5g yeast extract, 3g NaCl, 2g KCl and 6g agarose.

Ampicillin (Amp, Sigma)- This antibiotic was added to agar and broth in order to select for bacteria carrying the ampicillin resistance gene on a plasmid. A stock solution of ampicillin was made up at a concentration of 50mg/ml in dH₂O. The solution was filter sterilised and stored at -20°C. It was added to broth and agar to a final concentration of 60 μ g/ml.

Chloramphenicol (CA, Sigma)- This antibiotic was added to agar and broth to select for bacteria carrying the chloramphenicol resistance gene. A stock solution is made to a concentration of 34mg/ml in ethanol and stored at -20°C. Use at a final concentration of 170 μ g/ml.

5-Bromo-4-Chloro-3-Indolyl β -D-Galacopyranoside (X-gal) (Sigma)- X-gal is the substrate for β -galactosidase which when metabolised turns blue. It is added to agar to allow for blue/white colour selection of colonies carrying the plasmid with the β -galactosidase gene. If the plasmid is recombinant, the β -galactosidase gene will be disrupted and the colonies will turn white as the X-gal will not be broken down.

Blue colonies will have the intact plasmid. Stock solution was made at a concentration of 20mg/ml in dimethylformamide (DMF, Sigma). It was stored protected from light at -20°C and used at a final concentration of 40µg/ml.

Isopropyl β-D-Thiogalactopyranoside (IPTG, Sigma) - IPTG is a derepressor of the Lac operon. It was added to agar to induce the β-galactosidase gene and thus in conjunction with the X-gal could be used for the blue/white selection of colonies. Stock solutions were made at a concentration of 100mM. It was stored at -20°C and used at a final concentration of 0.5mM.

2.1.3 Bacterial Glycerol Stocks

10mls of bacterial culture were grown in a 37°C shaker for 4 hours, until opaque. Cells were pelleted at 3500rpm for 10 minutes and then re-suspended in 2mls of 15% glycerol in L-Broth. They were then frozen in liquid nitrogen and stored at -70°C.

2.1.4 Preparing Bacterial Competent Cells

Bacterial cells were streaked out from glycerol stocks onto an L-Agar plate and incubated at 37°C overnight. 10mls of L-Broth were inoculated with a single colony and grown overnight, shaking in a 37°C incubator. The overnight culture was then used to inoculate a further 500mls of L-Broth and left to shake in the 37°C incubator until they reached log phase ($OD_{590}=0.7$).

Cells were pelleted by spinning at 5000rpm for 5 minutes, and washed with 250mls of cold 0.1M $MgCl_2$ and pelleted again. They were re-suspended in 250mls of cold 0.1M $CaCl_2$ and incubated on ice for 20 minutes. Cells were harvested and re-suspended in 42.5mls 0.1M $CaCl_2$ and 7.5mls sterile glycerol. Aliquots were frozen in liquid nitrogen and stored at -70°C.

2.1.5 Transformation of Competent Cells

Competent cells were thawed on ice for 15 minutes. 10µls of a plasmid were mixed with 100µls of cells and incubated on ice for 30 minutes. Cells were heat shocked at 37°C for 2 minutes and then incubated on ice for a further 15 minutes. The transformation mixture was then plated on pre-warmed L-Agar plates containing the appropriate selection.

2.1.6 Small Scale Preparation of Plasmid DNA

Manual Minipreping

When large numbers of transformed colonies had to be screened for recombinant plasmids, the alkaline lysis method of plasmid preparations was used according to Molecular Cloning: A Laboratory Manual.

2mls of L-Broth containing the appropriate antibiotic was inoculated with a single bacterial colony and grown overnight at 37°C with vigorous shaking. Cells were spun down in a microfuge and the growth medium was removed, leaving the bacterial pellet as dry as possible.

The bacterial pellet was then re-suspended in 100µls of solution I (50mM glucose, 25mM 2-amino-2(hydromethyl)-1,3-propanediol (Tris).Cl (pH 8.0), 10mM ethylenediaminetetra-acetic acid disodium salt (EDTA) (pH 8.0)) followed by 200µls of solution II (0.2M NaOH, 1% sodium dodecyl sulphate (SDS)) and the tube inverted several times to mix the contents. The tubes were stored on ice for 5 minutes. 150µls of solution III (3M acetate, 5M potassium ions) was added and the tubes mixed to disperse the lysate through solution III. Tubes were stored on ice for 5 minutes and then spun at 12,000g for 5 minutes. A phenol:chloroform extraction was carried out on the supernatant to remove any salts, proteins and solvents from the DNA. The double-stranded DNA was precipitated with 2 volumes of ethanol, and the DNA pellet was then washed with 1ml 70% ethanol, and the pellet allowed to air-dry. The DNA pellet was then dissolved in 50µls of 10mM Tris.Cl (pH8.0), 1mM EDTA (pH 8.0) (TE) containing DNase -free RNase (20µg/ml).

Qiaprep Spin Plasmid Kit

For increased purity of the DNA required for sequencing, a Qiagen kit was used according to the manufacturer's instructions.

2.1.7 Large Scale Preparation of Plasmid DNA

For large pure quantities of DNA, Qiagen maxiprep (500 Tips) kits were used, according to the manufacturer's instructions.

2.2 Yeast Two Hybrid Work

2.2.1 Strain Used

The strain of yeast used for the yeast two hybrid (Y2H) system was Y190, whose genotype is: MATa, gal4, gal80, his3, trp1-901, ade2-101, ura3-52, leu2-3, -112

URA3::GAL-lacZ, LYS2::GAL(UAS)-HIS3,cyh.[†] This strain therefore has lacZ and the yeast biosynthetic gene HIS3 as the reporter genes under the control of Gal4-upstream activator sequence (UAS).

2.2.2 Vectors

The vectors used in the Y2H were the Gal4p DNA-binding domain vector pAS1-CYH2 (gift from Dr. Elledge, Texas) and the Gal4p activation domain vector pGAD10 (Clontech). The genes of interest had to be subcloned in frame to produce fusion proteins with either the DNA-binding domain or the activation domain.

2.2.3 cDNA Libraries

The cDNA library used in the screen was the Clontech human foetal kidney library, made up as a Gal4p transcription activation domain fusion library. A cDNA library from the mesonephric cell line, M15, was made in house together with Dr. Michael Lodomery by following the HybriZAP Two-Hybrid cDNA Gigapack Cloning Kit from Stratagene.

2.2.4 Yeast Transformations

A preculture of Y190 was made by inoculating 10mls of YPD (10g yeast extract, 20g peptone, 20g D-glucose per litre, autoclave) supplemented with 200µls of adenine (2mg/ml) with a single colony of the yeast. The culture was placed at 30°C overnight.

100µls of the preculture was then inoculated into 100mls YPD supplemented with 1ml of adenine (2mg/ml). It was allowed to grow at 30°C until it reached an OD₆₀₀ of 0.7. The yeast was then spun down at 2000rpm for 5 minutes and washed in 20mls sterile water. This was repeated again and after the second wash the yeast were re-suspended in 1ml LiAcTE (100mM LiOAc, 10mM Tris pH 8, 1mM EDTA. Autoclave) and transferred to an eppendorf. They were then spun for 20s and re-suspend in 1ml LiAcTE, so as to make yeast competent to take up the vectors to be transformed. This step was then repeated. The yeast were re-suspended to a final volume of 1ml in LiAcTE and kept on ice.

Meanwhile the transformation material was prepared by adding 0.5µg of the plasmid DNA with 5µg of library DNA, together with 50µg of salmon sperm DNA. 50µls of the prepared yeast was added to each transformation tube as well as 300µls 40% polyethylene glycol (PEG, MW 3350, dissolved in 1 x LiAcTE) and the mixture was vortexed. The tubes were incubated at 30°C for 30 minutes followed by heat shock

at 42°C for 20 minutes. The tubes were then spun again and the PEG removed, leaving no trace of it. The yeast were re-suspended in 1ml of YPD and incubated at 30°C for 1-2 hours. The yeast were then plated out onto AH and 50mM AT plates (see below) and incubated at 30°C for up to a week.

Yeast plates are made as follows:

20g/l glucose, 6.7g/l yeast nitrogen base without amino acids, 20g/l bacteriological agar (Oxoid), 800mg adenine. Amino acids and additives were added as required (see below) and the media was then autoclaved.

AH plates: Add histidine at 20mg/l. These plates select for yeast that have both vectors transformed and thus can synthesise their own leucine and tryptophan.

50mM 3-AT: Add 3-amino-1,2,4-triazole (3-AT, Sigma). Yeast that have been transfected will express the two proteins encoded by the cDNA cloned into the vectors described in section 2.2.2. If the two proteins interact, they will reconstitute the Gal4 transcription factor which will activate transcription of the His3 gene, and therefore will grow on plates that have not been supplemented with histidine. To ensure there is no background level of histidine 3-AT is added to the plates otherwise yeast which do not have the two interacting proteins would grow. 3-AT removes basal levels of histidine and so allows for the selection of yeast carrying two interacting proteins.

2.2.5 β -Galactosidase Yeast Filter Lifts Assay

The yeast strain used carries a β -galactosidase gene as a second reporter and this reporter can therefore be used to determine whether an interaction between two proteins of interest is occurring. It was therefore used as a way of eliminating false positives during a library screen as well as trying to determine whether deletion constructs of one particular protein can still bind to its partner protein in the same way that the wild type protein could. The clones of interest, including positive and negative controls were taken from AH plates and streaked onto fresh AH plates. Once the colonies had grown a filter lift was carried out by transferring the colonies onto a filter paper (Whatmann qualitative grade 1). The filter paper was then dipped into liquid nitrogen to permeabilise the yeast, and the filter was allowed to thaw. It was then placed, yeast side up, on top of another filter paper which was soaked in 0.18ml 10x buffer Z (0.6M Na₂HPO₄, 0.4M NaH₂PO₄, 0.1M KCl, 0.01M MgSO₄, 0.5M β -mercaptoethanol) and 30 μ l X-gal (20mg/ml in DMF). The filters were incubated at 30°C until a blue colour developed.

2.2.6 Liquid β -Galactosidase Assay

This assay can be used to compare the different strengths of the interactions between proteins (Estojak et al., 1995). The yeast transformed with the plasmids of interest were grown in 10mls of AH media and grown to OD₆₀₀ of 0.7. The yeast were then spun down at 2000rpm for 5 minutes and re-suspended in 1ml of sterile water. The yeast were spun down again and re-suspended in 250 μ ls Tris pH 7.5/0.05% Triton X-100 and frozen in liquid nitrogen and stored at -70°C until required for further use. 200 μ ls of this yeast extract was then added to 0.8mls of 1x buffer Z (see section 2.2.5) and 0.2ml of ONPG (4mg/ml). They were incubated at 30°C until a yellow colour developed, with the time of reaction being noted. The reaction was then terminated by adding 0.3ml of 1M Na₂CO₃, and it was placed on ice. The debris were spun down and the OD₄₂₀ was measured. In the meantime a protein concentration reading from 5 μ ls of the yeast extract was taken using the BioRad protein assay (see section 2.6.4). The β -galactosidase activity could then be calculated using the following formula:

$$\text{Activity} = (1000 \times \text{OD}_{420}) / (\text{volume of yeast used/ml} \times \text{time for colour to develop/minutes} \times \text{protein concentration/mg/ml})$$

2.2.7 Recovery Of Plasmids From Yeast Into *E. coli*

The plasmid of interest could be recovered and re-tested against other bait plasmids used as negative controls, or it could be sequenced. To recover the plasmid, the yeast had to be streaked onto AHT plates (AH plates containing tryptophan at 20mg/l, which was added after autoclaving). As the vector containing the tryptophan selectable marker is no longer required the yeast start to lose the vector at a low frequency when grown at 30°C. A single colony that had grown on these plates was then grown on AHT-cyclohexamide plates (2.5 μ g/ml of cyclohexamide dissolved in water and filter sterilised). The Y190 strain is resistant to cyclohexamide due to a mutation in the CYH2 gene, but the pAS1-CYH2 plasmid contains a wild type CYH2 gene thus causing the yeast to be sensitive. Thus when the yeast are grown on this medium, only yeast which have lost the pAS1-CYH2 plasmid can survive. To ensure that this plasmid has been lost, a β -galactosidase yeast filter lift can be done, and there should be no blue colour developing.

The remaining pGAD10 plasmid was then extracted from the yeast by adding a loop of yeast from the streaked yeast on the AHT-cyclohexamide into an eppendorf containing plasmid release buffer (2% Triton X-100, 1% SDS, 0.1M NaCl, 0.01M Tris pH 8, 0.001M EDTA), 300 μ ls glass beads (Sigma), 100 μ ls phenol and 100 μ ls

chloroform. The yeast were vortex well for 2 minutes and spun for 5 minutes. A chloroform extraction was then carried out on the supernatant. The DNA was then precipitated (see section 2.3.8) and the plasmid could then be transformed into bacteria (see section 2.1.5).

2.3 DNA Manipulations

2.3.1 Restriction Endonuclease Digestion of DNA

DNA was digested with the appropriate endonuclease using the specified buffer, according to the manufacturers. The enzyme volume never exceeded 10% of the final reaction volume as the glycerol in the enzyme solution can inhibit the endonuclease. The digests were typically carried out for 2 hours at 37°C, except were specified by the manufacturers. If a double digest was to be carried out, then only if both enzymes required the same buffer and temperature where they used simultaneously. Otherwise, the enzyme requiring the lowest salt concentration was used for the digestion first, and after a couple of hours the second buffer and enzyme were added and allowed to digest for another two hours.

2.3.2 DNA Gel Electrophoresis

DNA molecules can be separated according to size using agarose gel electrophoresis. In general a 1% gel was run, but if smaller fragments were to be resolved, then a higher percentage gel was used. The gels were made in TBE buffer (20 X TBE: 1M Tris.Cl, pH 8.0; 20mM EDTA; 1M boric acid, pH 8.3) unless the fragments were to be purified from the gel in which case a low melting point agarose (Flowgen) was used and this was made using TAE buffer (20 X TAE: 0.8M Tris.Cl pH 8.0; 20mM EDTA; 0.4M acetic acid). The DNA to be run on the gel was first mixed with 10x loading buffer (20% ficoll; 100mM EDTA; Orange G (Sigma)) to one tenth of the final volume. A DNA ladder (λ DNA digest with Hind III together with ϕ X174 digested with Hae III) was also run on the gel for size markers. After electrophoresis the DNA fragments could be visualised on a UV transilluminator as the gel had been stained with ethidium bromide (0.5 μ g/ μ l).

2.3.3 Dephosphorylation of Linear Plasmid DNA

If subcloning is carried out with an enzyme that leaves compatible ends on the cut plasmid, the cut plasmid DNA has to be dephosphorylated to prevent the ends from re-annealing. Thus after the digest, 1 μ l of calf intestinal alkaline phosphatase (CIP) and its corresponding buffer were added and placed at 37°C for an hour. To

inactivate the CIP, the reaction was shifted to 75°C for 10 minutes. Before ligation, the plasmid had to be cleaned up to remove all the enzymes, and this was done by using a GeneClean Kit (Bio101).

2.3.4 Genecleaning

The GeneClean II Kit (Bio101) is used to remove and purify DNA from virtually any type or grade of agarose and to eliminate impurities such as RNA, proteins and also primers from PCR reactions. This procedure was carried out as instructed by the manufacturers.

2.3.5 Ligation of DNA Molecules

To construct recombinant plasmids with cohesive termini, the plasmid which had been digested, dephosphorylated and genecleaned (see sections 2.3.1, 2.3.3 and 2.3.4) was mixed with the DNA to be subcloned at equal molar ratios. If blunt end ligation was to be carried out, then the plasmid and DNA were mixed such that there was always an excess of the insert. 1µl of T4 DNA ligase was added to the plasmid and insert mix, together with 1µl of 10x ligase buffer. The reaction volume was made up to 10µls with H₂O and placed in a 16°C incubator overnight.

2.3.6 Phenol:Chloroform Extraction

This is used to purify DNA or RNA from salts, protein or solvents. Before use the phenol needs to be equilibrated with 100mM Tris pH 7.5 to ensure that when the DNA is purified it goes into the aqueous phase instead of into the phenol. The chloroform needs to be mixed in a 24:1 ratio with isoamyl alcohol to stop the chloroform from frothing. The phenol and chloroform can then be mixed in a 1:1 ratio and kept in the dark. To purify the DNA or RNA, an equal volume of phenol:chloroform was added, mixed by vortexing and centrifuged at 12,000rpm for 2 minutes. The aqueous supernatant was then transferred to a fresh tube for the nucleic acid to be precipitated.

2.3.7 Ethanol Precipitation

To concentrate DNA or RNA a tenth of the volume of 3M NaOAc was added to a DNA or RNA solution followed by 2-2.5 volumes of 100% EtOH (or alternatively 0.7 volumes of isopropanol can be used). The tubes were left at least 20 minutes on ice (or -70°C to increase yields) and then centrifuged at 12,000rpm for 40 minutes at 4°C. The pellet was washed in 70% EtOH and then spun again. The supernatant

was then removed and the nucleic acid pellet allowed to dry before re-suspending it in the desired volume of dH₂O.

2.3.8 Southern Blotting

Preparation of the Filters:

1µg of DNA was digested with the appropriate enzyme to release the insert from the vector, and the fragments were separated by agarose gel electrophoresis. The DNA was depurinated by soaking the gel in 0.25M HCl for 10 minutes (or until the bromophenol turns yellow). The DNA was then denatured by soaking the gel twice for 15 minutes in 0.5M NaOH, 1.5M NaCl, after which time the gel was blotted by capillarity onto Hybond N+ membrane. After the transfer, the DNA was crosslinked onto the membrane by UV crosslinking.

Radioactively Labelling a DNA Probe

The random prime DNA labelling kit (Boehringer Mannheim) was used to make the probe. The insert which was to be used as the probe was purified by Genecloning. 25-50ng of DNA was made up to 9µls with water, and the DNA denatured by heating to 100°C for 5 minutes and then cooled on ice.

2µls of reaction mixture (containing the random hexanucleotide primers and the reaction buffer) was added to the DNA followed by 1µl of each dATP, dGTP, dTTP; 1µl Klenow polymerase and 5µls of [α^{32} P] dCTP. The reaction was incubated at 37°C for 30 minutes. The probe was then purified from the reaction components by running the mixture through a Sephadex G-50 column (Pharmacia).

Hybridisation

The Hybond membrane was prehybridised in 5mls of 0.5M NaPPi, 1% SDS, 1mM EDTA, 0.1mg/ml heat denatured salmon sperm DNA and 5µg heat denatured vector (competitor) for at least an hour at 65°C.

The probe was denatured by heating to 100°C for 5 minutes and then cooled on ice. 250µls of prehybridisation solution was added to the probe and 100µls of this was then added to each hybridisation bottle containing the membranes. Hybridisation was carried out overnight at 65°C. The filters were then washed three times in 2 x SSC (20xSSC: 175g sodium chloride, 88.2g of sodium citrate per litre of water, pH 7.0), 0.2% NaPPi, 0.4% SDS at 65°C. They were then checked for background activity and if hot, washed once more before being placed inside plastic wallets, assuring there were no trapped air bubbles that would distort the signal, and then put down for autoradiography.

Stripping and Reusing the Filters

If the filter was to be re-probed, the filters could be re-used so long as they had not dried out. The filters were first stripped by soaking them in 0.4M NaOH at 50°C for 2 x 30 minutes. The stripping could then be checked by placing the filter down on the phosphoimager.

2.3.9 Screening a Mouse cDNA Library

A mouse 11-day embryo 5' stretch plus cDNA library from Clontech was screened using [³²P]-NAL probe as described in section 2.3.8. Approximately 1x10⁶ clones from the library were screened with the probe. Hybridisation of the filters was carried out at 65°C in 4x SSC, 0.4%SDS, 0.2%NaPPi, 100µg/ml salmon sperm DNA and 2x denhardts. Washes were carried out at 65°C in 2xSSC, 0.4%SDS and 0.2% NaPPi. Positive clones were subjected to secondary and tertiary screening. The vectors containing the Nal inserts were then recovered by using a Qiagen lambda prep kit, and finally the inserts were subcloned into Bluescript vector and sequenced.

2.3.10 Primer Production

Primers were ordered for synthesis by Genosys.

2.3.11 Polymerase Chain Reaction (PCR)

PCRs were carried out in 500µls eppendorf tubes in a total volume of 50µls.

If the PCR product was to be cloned and used further, then the proof reading DNA polymerase Deep Vent (New England Biolabs) was used. If the PCR was used to only detect the presence of a stretch of DNA then the cheaper polymerase, AmpliTaq, was used instead. PCR reactions were set up to get a final concentration of the following: a 1x concentration of the polymerase buffer, 2mM MgSO₄, 200µM of each dNTP, 1µM each of a forward and reverse primer, 25-100ng of DNA template and 1 unit of the polymerase. The PCR reactions were carried out in a Hybaid Omnigene machine using the appropriate conditions for each PCR reaction.

2.3.12 ABI Sequencing

All sequencing reactions were carried with the automated sequencer and the Dye Terminator Sequencing Kit (Perkin Elmer, Applied Biosystems).

The reactions were set up as follows:

Terminator Ready Reaction Mix - 8µls
Pure DNA Template (Qiagen prepared)- 1µg
Primer - 3.2µls/µM or 3.2pmols
H₂O - up to 20µls
Overlay with mineral oil to avoid evaporation.

The PCRs were set up in a Hybaid Omnigene, under the following conditions:

94°C for 30 seconds
50°C for 15 seconds
60°C for 4 minutes
The reaction was allowed to proceed for 25 cycles.

The DNA was then precipitated by adding the PCR mixture to 2µls 3M NaOAc (pH 7.2) and 50µls EtOH. It was left on ice for 15 minutes and spun for 20 minutes before the precipitate was washed with 250µls of 70% EtOH. The pellet was air dried. At this point the pellet was handed over to the technical assistant for loading on to the gel, and formatting the output. It is to be noted that the later sequences were obtained by using rhodamine instead of the terminator ready reaction mix, as the system was upgraded.

2.4 RNA Manipulations

2.4.1 Northern Blotting

RNA Preparation From Tissue

The organs required for RNA extraction were dissected out of an adult mouse and stored on ice. 3mls of 3M LiCl/ 6M Urea were added to the tissue which were immediately homogenised. The LiCl causes RNA precipitation whilst the urea denatures proteins to minimise RNase activity. The samples were then sonicated for 1 minute at the power setting of 22 to shear the genomic DNA, and the RNA was then left to precipitate overnight at 4°C. The RNA was then pellet by spinning at 12,000rpm at 4°C for 20 minutes and all traces of the LiCl/Urea removed. The pellet was re-suspended in 0.3mls 10mMTris/0.5%SDS, and a phenol/ chloroform extraction followed by chloroform extraction was carried out to purify the RNA. The RNA was then precipitated by adding 30µls NaOAc and 1ml 100% EtOH and left at -70°C overnight to increase yields. The RNA was washed with 70% EtOH and

precipitated again, and finally was re-suspended in 50-100µls DEPC-H₂O, according to the size of the pellet.

Gel Electrophoresis of RNA Samples

When running RNA out on a gel, care must be taken to ensure that RNases do not degrade the RNA. The gel tank was therefore treated overnight with DEPC-H₂O to remove all traces of RNase. The gel itself contains formaldehyde to inactivate any RNase, and was made by dissolving 1g agarose in 72mls dH₂O and 10mls 10x electrophoresis buffer (20mM 3-[N-morpholino]propanesulfonic acid (MOPS), 5mM NaOAc pH7.0, 1mM EDTA). When the solution had cooled to 60°C, 18mls formaldehyde and 50µg ethidium bromide were added and the gel cast. 15µg of RNA were added to 15µls sample buffer (50%formamide, 2.2M formaldehyde, 1x electrophoresis buffer) and then heated to 55°C for 5 minutes. 5µls loading dye were added to the samples before they were loaded onto the gel. The gel was then run overnight, to give better resolution of the bands, at 10V.

RNA Transfer

The RNA was transferred to a nitrocellulose filter so as to be able to do the hybridisation step. Transfer was done by capillarity by placing a piece of Whatmann 3MM paper on a glass plate, soaked in and dipping into transfer buffer (16xSSC, 2.2M formaldehyde) with any air bubbles removed using a glass rod. The gel was then placed over this, followed by a Hybond N+ filter, cut to the size of gel, soaked in transfer buffer. 2 more pieces of Whatmann paper, cut to size and soaked in transfer buffer were then placed on top of the filter, followed by several paper towels, ensuring there was no short circuiting of fluid. A weight was then placed over the stack of paper towels and transfer was allowed to proceed overnight.

Prehybridisation of the Filter

After transfer was complete, the filter was briefly washed in 2xSSC to remove residual formaldehyde, and then baked at 80°C for an hour. The filter was then prehybridised (50% formamide, 5xSSPE (3m NaCl, 0.3M NaOAc, 2mM EDTA) 1% SDS, 5x denhardt (50xdenhardts: 5g ficoll, 5g polyvinyl pyrrolidone, 5g BSA. Make up to 500mls with dH₂O and filter sterilise), 100µg/µl ssDNA) for at least 30 minutes, prior to the probe being added and allowed to hybridise at 42°C overnight.

Washing the Filter

The filter was then washed to remove any unbound probe, with 2x SSC/0.1% SDS for 2x 10 minutes, followed by 3x10 minutes of 0.2x SSC/ 0.1%SDS at 60°C. The

filter was then wrapped in Clingfilm and placed in a phosphoimage cassette overnight, before it was developed.

2.4.2 Reverse Transcriptase PCR (RT-PCR)

To detect the presence of certain RNA's in tissues, an RT-PCR can be carried out. RNA was first extracted from the tissue (see section 2.4.1 - RNA preparation from tissue) and 5µg of the RNA in a total of 5µls of water was denatured by heating to 90°C. After cooling on ice the following was added to the tube: 10µls of 5x RT-buffer, 5µls 10mM dNTPs, 1µl RNase inhibitor, 1µl of random hexamers, 27µls DEPC-H₂O, 1µl of reverse transcriptase. This was then placed at 37°C for an hour to allow for the cDNA to be synthesised. A standard PCR reaction (see section 2.3.11) can then be carried out using the RT-product as a template.

2.4.3 DIG Labelling An RNA Probe

RNA probes can be labelled with digoxigenin (DIG) and detected after homologous binding to endogenous RNA by enzyme-linked immunoassays using an anti-digoxigenin alkaline phosphatase conjugate (anti-DIG-AP). A subsequent enzyme-catalysed colour reaction with 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) and nitroblue tetrazolium salt (NBT) produces an insoluble blue precipitate, thus detecting the hybrid molecules.

The DNA to be transcribed was cloned into bluescript vector such that it had T3 and T7 promoters adjacent to the polylinker. The DNA template was linearised so that the RNA polymerase produced "run-off" transcripts, which were then purified using a gel extraction kit (Qiagen). The DIG-labelled probe was then synthesised according to the DIG RNA Labelling Kit from Boehringer Mannheim, where DIG-UTP is used as one of the nucleotides which is incorporated into the transcript, together with CTP, ATP, and GTP. Typically 10µg of DIG-RNA was produced from 1µg DNA. The newly labelled RNA was then precipitated and run on a 1% agarose gel to check the integrity.

2.5 *In Situ* Hybridisations

2.5.1 Whole Mount *In Situs*

This protocol taken from Hammond et al., 1998 was used to determine the gene expression pattern during embryonic development in mice. Embryos of the desired age were dissected in DEPC-treated PBS (dPBS) and they were then fixed immediately in 4% paraformaldehyde (PFA, Sigma) overnight and then dehydrated

by placing in 100% methanol. The embryos could then be stored in methanol at -20°C.

Pre-Hybridisation

The embryos were first re-hydrated after being stored in methanol as follows:

75% methanol: 25% dPBS: 0.1% Triton X-100 – 5 minutes on ice

50% methanol: 50% dPBS: 0.1% Triton X-100 – 5 minutes on ice

25% methanol: 75% dPBS: 0.1% Triton X-100 – 5 minutes on ice

The solution was then replaced with dPBS/0.1% Triton X-100 for 3 x 15 minutes to permeabilise the cells.

Embryos were then treated with proteinase K (10 µg/µl; Boehringer Mannheim) in dPBS/0.1% Triton X-100 to remove the proteins allowing the probe to penetrate. Treatment varied from 15-25 minutes according to the age of the embryos. The embryos then had to be fixed by treating with 4% PFA for 45 minutes on ice.

The fix was then replaced with prehybridisation solution (25mls de-ionised formamide (Sigma); 12.5mls 20xSSC; 1g blocking powder (Boehringer Mannheim); 500 µls 10% Triton X-100; 50 µg/µls heparin (Sigma); 1mg/ml yeast-RNA (Sigma); 0.5mls 0.5M EDTA; 2.5mls 10% 3-[(3-Cholamidopropyl)dimethylammonio]-1-propane-sulfonate (CHAPS, Sigma) for 2x5 minutes at room temperature. The embryos were prehybridised for 1 hour at 65°C, and the solution was then replaced with fresh prehybridisation solution and left for a further 2-4 hours at 65°C, after which time the probe (see section 2.4.3), which had first been denatured at 80°C, was added and left to hybridise overnight at 65°C.

Post-Hybridisation

The next day the unbound probe was washed off as follows:

100% post-hybridisation solution for 5 minutes;

75% post-hybridisation; 25% 2xSSC for 10 minutes at room temperature;

50% post-hybridisation; 50% 2xSSC for 20 minutes at room temperature;

25% post-hybridisation; 75% 2xSSC for 30 minutes at room temperature.

Post-hybridisation solution: 50% formamide, 5xSSC, 0.1% Triton X-100, 0.5% CHAPS.

Further washes were carried out at 55°C for 2x30 minutes with 2xSSC/0.1%CHAPS, followed by one quick wash and then 2x30 minutes washes at 55°C with 0.2xSSC/0.1%CHAPS.

TNT (0.1% Triton X-100; 50mM Tris pH7.5; 150mMNaCl) was then added for 2x5 minutes at room temperature.

Signal Detection

The embryos then had to be blocked (0.1% Triton X-100; 50mM Tris pH 7.5; 150mMNaCl; 1g BSA; 15% sheep serum) for 3-4 hours at 4°C so that there was no background binding of the antibody. This solution was then replaced with fresh block containing the anti-DIG antibody (1/2000 dilution) and incubated at 4°C overnight.

The following day, the excess unbound antibody had to be removed by several washes: TNT/0.1%BSA for 4x1hour followed by 2x30 minutes with TNT.

The embryos were then placed in AP buffer (0.1M Tris pH 9.5; 0.05M MgCl₂; 0.1M NaCl) for 3x10 minutes, to ensure they were at the correct pH for the phosphatase to act. The colour reaction was carried out by adding 45µls NBT (dissolve 0.5g NBT in 10mls of 70% DMF, store at 4°C) and 35µls BCIP (dissolve 0.5g of BCIP in 10mls of 100% DMF, store at 4°C) to fresh AP buffer and this was added to the embryos and incubated in the dark for 1-3 hours to allow for the colour reaction to occur. Once the signal had appeared, the embryos were fixed in 4% PFA overnight at 4°C after which time the PFA is replaced by PBS.

2.5.2 *In Situ* Hybridisation For Sections

DIG-labelled *in situs* can also be carried out on sections through mouse embryos for a more precise internal examination of where a gene is expressed.

Section Preparation

Dissected embryos of the required age were fixed in fish fix (4% PFA in fish fix buffer: 4% sucrose in non-DEPC treated PBS) at 4°C overnight. The embryos were then washed for 3x5 minutes in fish fix buffer. They were embedded in 1.5% agarose, 5% sucrose in PBS. Care was taken to ensure the embedding mixture was not too hot or starting to set already. After the blocks containing the embryos had hardened they were trimmed to size and transferred to 30% sucrose in PBS and kept at 4°C until the blocks had sunk- usually overnight. The blocks were then frozen in dry ice by placing on top of a piece of aluminium foil which in turn was on top of a

metal rack which was covered by dry ice. At this stage the blocks could be kept at -70°C for up to a year.

For sectioning the blocks, a cryostat machine was used. The embryo blocks were fixed onto the chuck using OCT. The chamber temperature was set to -24°C and sections cut to 10-16µm thick. As sections were cut they are transferred onto 3-aminopropyl-triethoxy silane (TESPA) slides - see below. The sections were allowed to air dry for up to 2 hours and then placed in a box with silica gel and were stored at -20°C until required.

TESPA slides

To avoid sections from floating off slides, the slides can be TESPAs coated as follows:

Slides were dipped in 10% HCl in 70% EtOH for 20 seconds; followed by distilled H₂O for 20 seconds and finally 100% acetone for 20 seconds. The slides were then allowed to air dry. For the second stage of coating, the slides were dipped into 2% TESPAs in acetone for 20 seconds followed by 100% acetone for 20 seconds and then into fresh acetone for another 20 seconds. The slides were allowed to air dry again before use.

Hybridisation

Cut sections had to be defrosted for an hour. The DIG-labelled probe (see section 2.4.3) was diluted 1 in 200 in hybridisation buffer (Hyb buffer: 1x salt (10x salt: 114g NaCl, 14.04g Tris HCl, 1.34g Tris base, 7.8g NaH₂PO₄·H₂O, 7.1g Na₂HPO₄, 100ml 0.5M EDTA in a total of 1000mls H₂O, pH 7.5; autoclave) 50% de-ionised formamide, 10% dextran sulphate, 1mg/ml yeast RNA, 1x denhardts). The mix was denatured at 70°C for 5 minutes before 100µl of the probe mix was added to each slide. A cover slip was added and the hybridisation allowed to proceed overnight in a Hybaid slide machine at 65°C.

Post Hybridisation Washes

Slides were transferred to a coplin jar and washed in solution A (1xSSC, 50% formamide, 0.1% Triton X-100) at 65°C for 15 minutes. Coverslips were removed and the slides were washed for 2x30 minutes at 65°C in buffer A. They were then rinsed at room temperature for 2x30 minutes in Tris buffered saline with Tween 20 (TBST).

Blocking and Antibody Staining

Slides were blocked in 10% heat inactivated sheep serum in 1xTBST for 1 hour at RT. The slides were then dried around the edges before applying 100µl of 1 in 2000 dilution anti-DIG AP Fab fragments in 10% heat inactivated sheep serum in 1xTBST. They were incubated overnight at 4°C.

Post Antibody Washes And Staining Reaction

5x20 minutes washes in TBST at RT, removing the coverslips after the first wash. 2x10 minutes washes in 1xNTMT (100mM NaCl, 100mM Tris HCl pH 9.5, 50mM MgCl₂ and 0.1% Triton X-100. The solution should be made fresh on the day from stock solutions). The slides were then stained in the dark with 4.5µl NBT and 3.5µl BCIP diluted in 1ml NTMT. A total of 30mls staining solution was needed per 10 slides. The staining was checked after a few hours.

Post Staining and Mounting

The staining reaction was stopped by washing twice in water. The stain was then fixed in 4% PFA/0.1% glutaraldehyde for 20 minutes. Sections were then dehydrated through an alcohol series. The slides were counterstained with 0.1% eosin in 95% EtOH for 20 seconds, and rinsed twice in 95% EtOH. They were transferred to histoclear for 2x5 minutes and then mounted in DPX.

2.6 Protein Work

2.6.1 In Vitro Transcription and Translation of Proteins

The Promega TNT coupled reticulocyte lysate system was used to make *in vitro* protein starting from a DNA template. By using [³⁵S]-methionine in the reaction the protein could be radioactively labelled for detection (see below, section 2.6.8). The protocol was followed according to the manufacturers' instructions.

2.6.2 Expressing pET6H-WT1 Proteins in Bacteria

Proteins can be expressed with 6 histidine molecules attached to them by using the pET6H vector. The 6 histidines then allow the protein to be purified using a nickel column.

A preculture of BL21 bacterial cells transformed with the pET6H-WT1 vector were grown overnight in LB containing ampicillin. The next day a 500mls culture was grown from the preculture and left until they reached an OD₆₀₀=0.7 (at this stage 1ml of the culture was removed for analysis). The rest of the culture was induced at 30°

C with 100mM isopropyl β -D- thiogalactoside (IPTG). Samples were then removed again after 30 minutes; 1 hour; 2 hours and 3 hours. To test whether the induction had worked, the 1ml bacterial samples were re-suspended in cracking buffer (10mM NaPPi pH 7.2, 1% SDS, 8M Urea, 0.1% Triton X-100, 1% β -mercaptoethanol, 0.15% bromophenol blue) and run on an SDS-PAGE (see section 2.6.7) to see if there was a strongly induced protein, corresponding to 6H-WT1.

Once the optimum expression time had been found from the time course experiment done, the solubility of the protein could be tested. The bacteria, which had been transformed with pET6H-WT1, were grown until they reached an $OD_{600}=0.7$. They were then induced with IPTG for 2-3 hours, after which time the bacteria were spun down. The pellet was then re-suspended in 100mls buffer A (10mM imidazole, 50mM NaCl, 10mM Tris.Cl pH7.9, 10% glycerol, 0.1% Triton X100, and one tablet complete protease inhibitors per 50mls (Boehringer Mannheim)). The bacteria were freeze thawed and then sonicated using a Lucas Soniphore sonicator, power 5 ensuring the sample was kept on ice. The solution was then left rotating at 4°C for 30 minutes before being centrifuged. The supernatant, or soluble fraction, was mixed with 2x sample buffer and the pellet was re-suspended in cracking buffer before being run on an SDS-PAGE gel to see in which fraction the protein is found in.

2.6.3 Expressing GST-Fusion Proteins in Bacteria

A protein of interest can be overexpressed in bacteria and then purified from the rest of the bacterial extract by using a glutathione S-transferase (GST) tag on the protein. The pGEX vector (Pharmacia Biotech) is an inducible expression vector which fuses the GST to your protein of interest, and this allows for easy purification of the protein because the GST tag can bind to a glutathione sepharose 4B affinity matrix (Pharmacia Biotech), whilst the rest of the bacterial extract gets washed through it.

The vector was transformed into BL21-pLysS bacteria (the pLysS expresses a lysozyme which aids in breaking down the bacterial cell wall, and this vector contains the CA resistance gene, see section 2.1.2). One colony was then grown overnight in 20mls LB/Amp/CA to select for both the pLysS and pGEX vectors. The next day, 400mls LB/Amp/CA were inoculated with 20mls of the preculture and allowed to grow at 30°C to reach an OD_{590} of 0.7 (at this stage 1ml of the culture was removed for analysis). The rest of the culture was induced at 25°C with 100mM isopropyl β -D- thiogalactoside (IPTG) for 2-3 hours (after which time 1ml was removed for analysis). To test whether the induction had worked, the 1ml bacterial

samples were re-suspended in cracking buffer (10mM NaPPi pH 7.2, 1% SDS, 8M Urea, 0.1% Triton X-100, 1% β -mercaptoethanol, 0.15% bromophenol blue) and run on an SDS-PAGE to see if there was a strongly induced protein, corresponding to the GST-fusion protein of interest.

2.6.4 Purification of the GST-Fusion Protein

The induced bacterial culture was pelleted by centrifuging at 6000rpm for 10 minutes, and the bacterial pellet re-suspended in 40mls of GST-binding buffer (50mM Tris pH 7.5, 100mM NaCl, 10% glycerol, 1% Triton X-100, 1 tablet of complete protein inhibitors per 50mls (Boehringer Mannheim)). The bacteria were freeze thawed and then sonicated using a Lucas Soniphore sonicator, power 5 ensuring the sample was kept on ice. The solution was then left to rotate at 4°C for 30 minutes to allow the proteins to solubilise, after which time they were centrifuged at 12,000rpm for 30 minutes to pellet insoluble cellular material. The supernatant, which contained the soluble proteins, was collected. The pellet was re-suspended in cracking buffer and compared to the soluble fraction on an SDS-PAGE to establish whether the protein of interest was soluble. Solubility of the protein could be optimised by altering the inducing conditions, such as the IPTG concentration, the inducing temperature or the length of time of induction. The GST-fusion protein was then purified from the rest of the fraction by using the prepared affinity matrix. The affinity matrix was prepared as follows: for the required bed volume of 200 μ ls of glutathione sepharose 4B, 266 μ ls were used as the matrix was supplied as a 75% slurry. The slurry was sedimented and washed in PBS to remove the EtOH. The slurry was then re-suspended in 200 μ ls PBS to form a 50% slurry. This slurry was then mixed with the soluble bacterial sample and left to spin for 30 minutes at 4°C. The mixture was then placed into a column, where any protein not bound to the slurry run through. The slurry was then washed 3 times with GST-binding buffer to remove any non-specific binding proteins. The GST-fusion protein could then be eluted with 200 μ ls of glutathione elution buffer, and a sample run on an SDS-PAGE to see how pure the protein of interest was.

2.6.5 BioRad Protein Assay

Protein concentrations can be estimated by using the BioRad protein assay. An aliquot of protein was diluted in 800 μ ls water and 200 μ ls of the BioRad solution was added and the solution mixed. It was allowed to stand for 10 minutes at RT for the colour to develop and an OD₅₉₅ taken. To calculate the concentration of the protein, a standard curve had to be produced using known concentrations of BSA.

The graph produced could then be used to read off the concentration of the protein using its OD value.

2.6.6 *In Vitro* Binding Assays

In vitro binding assays can be carried out to see if two proteins interact. 5µg of the GST-fusion protein of interest (see section 2.6.4) was mixed with *in vitro* translated (IVT) (see section 2.6.1) protein together with 200µls of low salt association buffer (LSAB) (100mM NaCl, 100mM Tris.Cl pH 8, 0.1% NP40). For more specific binding the salt concentration could be increased to 250mM and was then called medium salt association buffer, or the salt concentration could be increased to 500mM, in which case it was called high salt association buffer. The mixture was left on a rotor at 4°C for 2 hours. Then prepared glutathione sepharose 4B (see section 2.6.4) was added and left to rotate for another 30 minutes. The slurry was then precipitated and washed 6 times with the association buffer, to remove any protein not binding to the GST-fusion protein. After the final wash, 30µls of 2 x sample buffer (SB: 62.5mM Tris.Cl pH 6.8, 2% SDS, 2% β-mercaptoethanol, 10% glycerol, 0.1% bromophenol blue) was added to the slurry and after boiling the proteins were run on an SDS-PAGE followed by fluorography to detect the presence of the radioactive protein (see sections 2.6.7 and 2.6.8).

2.6.7 SDS-PAGE Electrophoresis

SDS-polyacrylamide gel electrophoresis (PAGE) was used to resolve proteins according to their molecular weight. The resolving gels were cast as shown in table 2.1.

The gel was assembled and the samples which had been mixed with 2 x SB were boiled for 5 minutes and loaded into the wells. Running buffer (25mM Tris, 250mM glycine pH 8.3 and 0.1% SDS) was added and the gel run at 150V.

The gel was then stained with Coomassie dye (0.1% Coomassie blue dye, 50% methanol, 10% acetic acid) for an hour before being placed into destain solution (5% methanol, 7% acetic acid) to remove background staining.

Solution components for resolving gel	Component volumes per 10ml of gel mould (ml)	
	8%	10%
H ₂ O	4.0	7.9
30% acrylamide mix	3.3	6.7
1.5M Tris (pH 8.8)	2.5	5.0
10% SDS	0.1	0.2
10% ammonium persulfate	0.1	0.2
TEMED	0.004	0.008

Solution components for a 5% stacking gel	Component volumes per 4ml and 8ml gel moulds (ml)	
	4ml	8ml
H ₂ O	2.7	5.5
30% acrylamide mix	0.67	1.3
1.0M Tris (pH 6.8)	0.5	1.0
10% SDS	0.04	0.08
10% ammonium persulfate	0.04	0.08
TEMED	0.004	0.008

Table 2.1: Solutions for preparing resolving (upper table) and stacking (lower table) gels for Tris-glycine SDS PAGE.

2.6.8 Detecting A Radioactive Protein

This is used when a radioactive ³⁵S-methionine protein needs to be detected. After the protein sample had been run on an SDS-PAGE, the gel needed to be fixed (50% EtOH, 10% acetic acid) for 2x30 minutes. The gel then had to be placed in Amplify (Amersham Science) for 15 minutes. This is an organic scintillant which converts the emitted energy of the isotope to visible light which can then be detected on X-ray film. The gel was then dried and placed overnight in a photographic cassette together with X-ray film.

2.6.9 Western Blotting

Transfer of Proteins Onto Hybond Filter

Western blotting is used to transfer proteins run on an SDS-PAGE, onto a Hybond filter. A semi-dry blotter was used which provided the anode and cathode plates.

The system was set up on the anode plate as follows: 2 Whatmann filters cut to the size of the gel were soaked in anode 1 (0.3M Tris, 10% methanol, pH 10.4), followed by another 2 pieces soaked in anode 2 (25mM Tris, 10% methanol, pH10.4). The Hybond filter, also cut to size, was equilibrated in anode 2 and placed on top of the Whatmann papers. The gel which had been soaked in anode 2 was then laid over the filter. 3 pieces of Whatmann soaked in cathode buffer (40mM glycine, 25mM Tris and 20% methanol, pH 9.4) were placed over the gel. The cathode plate was then put in place and the transfer carried out at 10V, 240mA for 30 minutes.

Detecting The Protein of Interest

Once the Western blot had been carried out the protein of interest could be identified using the appropriate antibody. First the nitro-cellulose membrane had to be blocked in blotto (5% fat-free milk in TBST) for 1 hour at room temperature and then incubated with the primary antibody overnight at 4°C. The unbound antibody was then washed off by carrying out three 15 minute washes in blotto. Secondary antibody, which recognises the primary antibody and is conjugated with horseradish peroxidase (HRP), was then added for an hour at room temperature, after which another three washes were carried out. Finally the filter was quickly washed in TBST and the antibodies detected using the ECL-Plus (Amersham), as described by the manufacturer.

Stripping The Filters For Re-Probing

If the filter had to be re-used so as to detect another protein, the filter first had to be stripped of the initial antibodies used. This was carried out by submerging the filters in stripping buffer (100mM β -mercaptoethanol, 2% SDS, 62.5mM Tris .HCl pH 6.7) and incubated at 50°C for 30 minutes. The filter was then washed in TBST for 2x10 minutes at RT. The filter was blocked with blotto and the immunodetection of the new protein of interest was then carried out as before.

2.7 Antibody Work

2.7.1 Antibody Production

A polyclonal antibody was to be raised against NAL in rabbits. A GST-NAL fusion protein was overexpressed in bacteria and purified (see section 2.6.3 and 2.6.4). A final concentration of 100 μ g of antigen was injected for each injection and subsequent boosts. Two rabbits were inoculated so as to increase the chances of obtaining a good antibody, and each rabbit was injected once followed by 3 booster

injections. The handling of the rabbits and the production of the sera was all carried out at the Scottish Antibody Production Unit (SAPU), Law Hospital, Carlisle, Lanarkshire ML8 5ES, Scotland, co-ordinated by Dr. Maggie Chambers.

2.7.2 Affinity Purification of NAL Antisera

A 10% SDS-PAGE (see section 2.6.7) was cast, using a preparative comb to run purified GST-NAL. A Western blot (section 2.6.9) was carried out and the filter was then stained with Ponceau red so that the strip of pure GST-NAL could be seen and cut out. This strip was then blocked in blotto, before being cut up into tiny fragments using a razor. 800µl TBS were then added to the fragments together with 200µl of the NAL antisera, and the tube were left to spin on a wheel for 2-3 hours at room temperature. The filter was then rinsed 3 times in TBST before washing 3 times with 1ml TBST for 30 minutes on the wheel. The filter was incubated with 200µl of 200mM glycine pH 2.5 containing 0.1% BSA for 30 minutes to elute the bound antibody. The supernatant was then collected and another 50µl of the glycine/BSA solution was added for another 10 minutes before that was collected and this was done twice over. The supernatant was then neutralised by adding 30µl 1M Tris (non-pHed). The purified antibody was finally dialysed using a 0.025µm filter for 10 minutes against TBS. The filters could be reused by washing them in TBST and then stored in TBS containing sodium azide.

2.7.3 Antibodies Used

Several antibodies have been used for Western blots and immunofluorescence. The antibodies used, together with some information about them are listed in the table below.

Antibody	Type of antibody	Supplier
Anti- IGC, 3C5	IgM Monoclonal	Bryan Turner
Anti-NAL, 474	Rabbit Polyclonal	Natalie Little
Anti-HA, 12CA5	Mouse Monoclonal	Boehringer-Mannheim
Anti-WT1, C19	Rabbit Polyclonal	Santa Cruz Biotechnology
Anti-WT1, H2	Mouse Monoclonal	DAKO
Mouse IgG	Mouse Monoclonal	Sigma

Table 2.2: List of antibodies used for Western blots and immunofluorescence.

2.8 Tissue Culture

2.8.1 Cell Lines

The cell lines used were the mesonephric cell line, M15 (Larsson et al., 1995), COS 7 and HeLa (European Collection of Cell Cultures, ECACC). They were grown at 37°C in a humidified, 5% CO₂ environment.

2.8.2 Medium

The medium used for all these cell lines was Dulbecco Modified Eagle Medium (DMEM, Gibco BRL), 500mls, supplemented with 5mls sodium pyruvate, 5mls penicillin and streptavidin antibiotic mix, 5mls glutamine and 50mls of heat inactivated foetal calf serum (FCS).

2.8.3 Splitting Cells

All cell lines used grew in an adherent manner and thus trypsin was used to remove cells from plates so that they could be used either in making nuclear extracts, to re-plate them at different densities or to freeze down as stocks. First the medium was aspirated off and cells rinsed twice with PBS. 1ml of trypsin was added to the cells and they were incubated at 37°C for 1 minute, after which time the cells started to come off the plate. The trypsin was then inhibited by adding the components of the FCS in the medium, and the cell suspension could then be used accordingly.

2.8.4 Freezing Cell Lines Down

The cell line to be frozen down for stock were washed with PBS and then trypsinised to remove the cells from the plate. 5mls of medium were added to stop the trypsinisation after the cells had come free from the plate, and the cells were collected and spun down at 1000rpm for 5 minutes. The medium was aspirated away and the cells re-suspended in 3mls of cold medium containing 10% DMSO. The cells were then aliquoted into vials and placed in a polystyrene box and placed at -70°C so that they froze down slowly after which time the vials were transferred to a liquid nitrogen tank for storage.

2.8.5 Growing Cell Lines Up From Frozen Stocks

A frozen down vial of cells was quickly thawed at 37°C. 5mls of medium was added and the cells were pelleted by centrifuging at 1000rpm for 5 minutes. The medium was then aspirated off to remove all traces of DMSO. The cells were then re-suspended in 5mls of medium and transferred to a flask and incubated at 37°C.

2.8.6 Transient Transfections

The cell line to be transiently transfected was grown on an appropriate sized plate and allowed to reach 50-80% confluency. To enable the DNA to be transfected, the plasmid had to be mixed with LipofectAmine (Gibco BRL) and the quantities vary according to the size of the culture plate used as follows:

Culture Plate/mm	Lipid and DNA dilution volume/ μ l	LipofectAmine volume/ μ l	Amount of DNA/ μ g	Transfection medium volume/ml
35	100	2-25	1-2	0.8
60	300	6-75	3-6	2.4
100	800	16-200	8-16	6.4

Table 2.3: Reagent quantities for transfection of DNA onto different sized culture plates.

The diluted DNA and lipid solutions were then gently mixed together and incubated at room temperature for 30 minutes to allow DNA-liposome complexes to form. Whilst these complexes form, the cells were washed in serum free medium, or opti-mem (Gibco BRL). The appropriate transfection medium volume, according to the table, was then added to the complexes and gently mixed, before overlaying on the cells. The cells were incubated at 37°C for 24 hours, after which time the transfection medium was replaced with complete medium, and cells were incubated for a further 24 hours. The cells could then be assayed for activity of the transfected gene.

2.8.7 β -Galactosidase Assay

As a positive control for transient transfections and to give an idea of the transfection efficiency, a β -galactosidase gene under the control of a CMV promoter was also transfected into the M15 cells. The transfection was carried out as for the other plasmids and 24 hours after transfection, the cells were rinsed in PBS and then fixed for 5 minutes at 4°C in 2% formaldehyde, 0.2% glutaraldehyde in PBS. The cells were then rinsed with PBS and then 1ml of X-gal reaction mixture (5mM potassium ferrocyanide, 5mM potassium ferricyanide, 2mM magnesium chloride and a final concentration 1mg/ml of X-gal in PBS, filter sterilised) was added to the chamber slides. The chamber slides were then incubated at 37°C for 1-18 hours to allow the blue colour to develop.

2.8.8 Co-Immunoprecipitations

Cross-Linking of Antibodies to Protein A Sepharose Beads

The antibody which was going to be used to immunoprecipitate a particular protein first had to be crosslinked to protein A sepharose beads. 1 μ g of the antibody was incubated with 50 μ ls of a 20% bead slurry per immunoprecipitation to be carried out. The mix was then made up to 1ml with IPB 100 (100mM sodium chloride, 10mM Tris and 1mM EDTA, made to pH 8) and left incubating at 4°C overnight. The beads were then spun down and washed twice in 0.2M sodium borate and then incubated for 30 minutes at room temperature in 0.2M sodium borate containing 20mM dimethylpimelidate pH >8.3. The beads were washed in 0.2M ethanolamine pH 8 and then incubated for 2 hours at room temperature in the same solution. The cross-linked beads were then washed twice in PBS, followed by two washes in 100mM glycine, pH 3 and then another further two washes in PBS. Finally the beads were re-suspended in 50 μ ls PBS.

Preparation of Nuclear Extracts

3 large plates of M15 cells were grown until confluent. Cells were then trypsinised (see section 2.8.3) and the cells pelleted by centrifuging for 5 minutes at 1200rpm, washed with PBS and re-pelleted. The cells were then re-suspended in 500 μ ls PBS and spun at 2000rpm for 3 minutes. The pellet was re-suspended in 5 times the packed cell volume of buffer A (10 X buffer A: 0.1M Hepes pH 7, 15mM MgCl₂, 0.1M KCl) containing 0.5 μ l/ml of 1M DTT and protein inhibitors. The solution was incubated on ice for 10 minutes and then spun at 2500rpm for 5 minutes. The supernatant was removed and the pellet was re-suspended in 2 times the packed cell volume of buffer A. The cell suspension was then homogenised on ice with a tight pestle for 10 strokes. The solution was spun for 10 minutes at 2500rpm, and the supernatant removed. The pellet was re-centrifuged at 13000rpm for 10 minutes to remove all the supernatant and the pellet that was left were the cell nuclei. These were then re-suspended in 1 ml of fresh buffer T (50mM Tris pH 8, 250mM NaCl, 1% Triton X-100, 0.2mM sodium vanadate, 5mM β -glycerophosphate and protein inhibitors). The nuclei were then spun at 13000rpm for 10 minutes and the supernatant was transferred to a fresh tube.

Carrying Out The Co-Immunoprecipitations

To carry out immunoprecipitations the cross-linked antibody was mixed with the nuclear extracts made, which had been previously pre-incubated with pre-blocking beads for one hour at 4°C to remove any proteins binding directly to the beads. The

mixture was incubated overnight at 4°C, after which time the beads were washed 6 times with IPB 100 containing 0.5% NP-40. The beads were finally re-suspended in 30µls of 2xSB, boiled and then run on 8% SDS-PAGE (see section 2.6.7) followed by Western blotting (see section 2.6.9) to detect the presence of the proteins of interest.

2.9 Fluorescence

2.9.1 Immunocytochemistry

M15 cells were grown on chamber slides (Nunc) until they had reached 80% confluency, after which time immunocytochemistry was carried out. The cells were rinsed twice in PBS and then fixed for 10 minutes at room temperature in an acetone: methanol (1:1) mix. The cells were blocked to reduce background in blocking solution (2% BSA, 0.2% Tween 20, 6.7% glycerol, 0.02% sodium azide, in PBS) for 1 hour at room temperature. The primary antibody was then added at the required concentration in blocking solution. 200µls of the primary antibody mix was used per 1 well chamber slide and overlaid with parafilm so as to minimise antibody usage. The cells were then rinsed 3 times in TBST for 5 minutes. The secondary antibody, which should recognise the heavy chain of the primary antibody and is tagged with a fluorescent label, was then added at a dilution of 1 in 100. Again 200µls were used per chamber slide which now had to be kept in the dark so that the fluorescence did not fade. The cells were incubated with the secondary antibody for 45 minutes after which time another 3 washes were carried out with TBST. The slides were then mounted with Vectashield containing 4',6 diamidino-2-phenylindole (DAPI, Vectra) to preserve the fluorescence and stain the nucleus. The fluorescence was then looked at under an Axioplan fluorescent microscope. Images were captured using a Digital Scientific Smartcapture software.

2.9.2 Immunohistochemistry

Immunohistochemistry was carried out on E=17.5d kidneys and testes. The organs were dissected out and immediately fixed in 4% PFA for 1-1.5 hours. They were then washed twice for 20 minutes in PBS, before placed overnight in a 30% sucrose solution made in PBS. The next day the organs were embedded in OCT and frozen in liquid nitrogen, and stored at -70°C. Sections, 10µm thick, were then cut onto TESPA slides using the cryostat (see section 2.5.2). The slides were left to dry at room temperature for 1 hour before the immunohistochemistry was carried out.

The sections on the slide were circled with wax, using a PAP pen, so as to use less quantities of antibodies. All subsequent steps were then carried out in a humidifying chamber to avoid the sections drying out. The sections were then blocked for an hour in blocking solution as was used in section 2.9.1. If the primary antibody to be used was monoclonal, then the sections had to be incubated with goat IgG Fab arms (1 in 25 dilution) for an hour to remove any background as the sections originated from mice. The sections were then washed 3x5 minutes in TBST. The primary antibody was then added, diluted accordingly to a total volume of 150µls in block. This was left for an hour before 3x5 minute washes with TBST were carried out. The secondary antibody, coupled to either FITC or Texas Red, was then added for 45 minutes and the slides were then kept in the dark from now on. The slides were then washed 3x5 minutes in TBST before being mounted with Vectashield containing 4',6 diamidino-2-phenylindole (DAPI, Vectra) to preserve the fluorescence and stain the nucleus. The fluorescence was then looked at under an Axioplan fluorescent microscope. Images were captured using a Digital Scientific *SmartCapture* software.

2.9.3 Fluorescent *In Situ* Hybridisation (FISH)

This was carried out in house by Muriel Lee on both human and mouse chromosomal spreads.

Chapter 3

Yeast Two Hybrid

Chapter 3: Yeast Two Hybrid System

In an attempt to identify proteins that interact with WT1, the yeast two hybrid (Y2H) system was to be used. The identification of proteins that interact with WT1 should help to elucidate its function and shed light on how WT1 works at the molecular level. Thus it is hoped that using this technique we will further our understanding of the role of WT1 in both development and tumourigenesis.

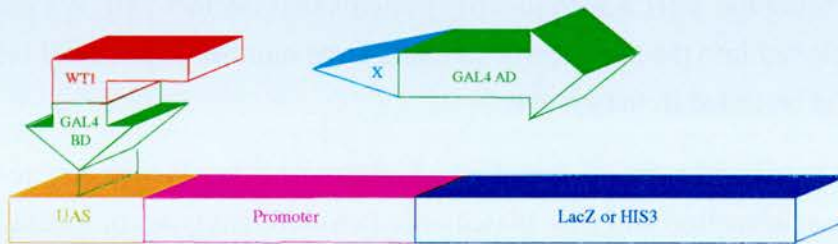
3.1 Introduction to the yeast two hybrid system

The Y2H system is a very powerful and sensitive tool used in the identification of protein-protein interactions. It is based on a yeast genetic assay in which the interaction of two proteins is measured by the functional restoration of a transcriptional activator (Fields and Song, 1989). There are a couple of variations on the system which use different strains of yeast, reporter genes and vectors - such as the system that uses the *Saccharomyces cerevisiae* Gal4p DNA binding and activation domains; or the one that uses the *Escherichia coli* LexA protein as the DNA-binding domain and the herpes virus VP16 protein as the activation domain; or even the one that uses the *E. coli* LexA protein as the DNA-binding domain and the *E. coli* B42 sequence as its activation domain. The system that was used, and therefore will be discussed here, is the one that uses the *Saccharomyces cerevisiae* Gal4p DNA binding and activation domains together with the *HIS3* and *lacZ* reporter genes.

The Gal4 transcription factor has two domains: a DNA-binding domain as well as a transcriptional activator domain, and the transcription factor is only active when both domains come together. The DNA-binding domain binds to an upstream activation sequence (UAS), which targets this domain to a promoter of a reporter gene, and the activator domain, once in a complex with the DNA-binding domain of Gal4, initiates transcription of the gene in question. The two Gal4 domains do not have to be in direct contact with each other and thus can be reconstituted through non-covalent interactions of two independent hybrid proteins which are fused to either domain. For the system to work, the hybrid proteins on their own should not activate transcription of the gene. Thus transcription is only brought about when the two hybrid proteins genuinely come together thus restoring the Gal4 transcription factor and causing expression of the gene. In order to know whether the two hybrid proteins interact, the genes transcribed are the reporter genes *lacZ* and *HIS3* (Durfee

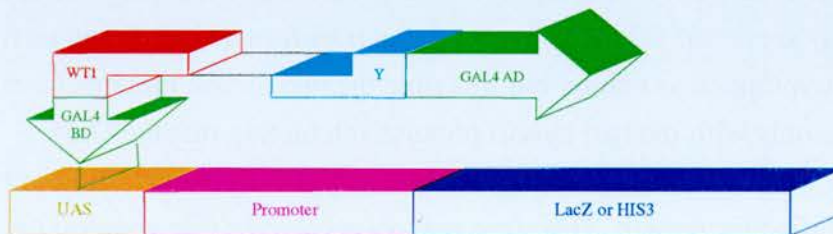
et al., 1993), figure 3.1. Any positive interactions can then be identified because the yeast carrying the two interacting proteins will grow on plates that lack histidine, and these yeast will also turn blue when a β -galactosidase assay is carried out.

A: No Expression Of Reporter Genes



The WT1-Gal4 DNA-Binding Domain (BD) binds to the upstream activator sequence (UAS), but because protein X does not bind to WT1 the Gal4 Activator Domain (AD) cannot localise to the Gal4 BD and so the Gal4 transcription factor is not reconstituted. The reporter genes are not expressed as the Gal4 AD cannot localise to the UAS on its own.

B: Expression Of Reporter Genes



If protein Y binds to WT1, the Gal4 BD and the Gal4 AD will be brought together, thus reconstituting the Gal4 transcription factor. The Gal4 BD will localise to the UAS and together with the Gal4 AD, they will express the reporter genes, *lacZ* and *His3*.

Figure 3.1: Schematic illustrating the basis of the yeast two hybrid system.

The first steps that need to be carried out when performing a yeast two hybrid screen, is the construction of the hybrid proteins. Usually the gene encoding the protein of interest, in this case *WT1*, is cloned into a vector that contains the DNA binding domain of Gal4 and this is done so as to form a fusion protein when the vector is expressed in yeast. This vector is referred to as the bait and it carries tryptophan as the selectable marker. The other vector, containing the Gal4 activator domain, is called the prey and carries leucine as the selectable marker. As the objective of using the Y2H was to identify proteins that interact with WT1, a cDNA library was cloned into the prey vector so that a large number of potential binding proteins could be tested in the system.

The bait and prey vectors are co-transformed into yeast using the lithium acetate method of transformation and then plated onto two different types of media. The first type is the AH medium, which contains added adenine and histidine (see materials and methods). This is because the yeast has a phenotype such that it is auxotrophic for adenine, tryptophan, leucine and histidine (see material and methods for the yeast Y190 genotype), and thus because the two vectors only provide leucine and tryptophan, the adenine and histidine have to be provided if the yeast are to grow. If transformed yeast grow on these AH plates then they must have taken up and expressed both vectors. The second type of medium is the 50mM 3'-aminotriazole (AT), which contains adenine and 50mM AT (see materials and methods) (Durfee et al., 1993). 3'-aminotriazole is added to the plates to dampen down any background levels of histidine which may be present due to the mutation in the *HIS3* gene in Y190 being leaky. When plated onto the 50mM AT medium, yeast can only survive if they have both vectors transformed into them, providing the leucine and tryptophan as before, but this time the two hybrid proteins have to interact. It is only with the two hybrid proteins interacting that the Gal4 transcription factor is reconstituted and the *HIS3* gene expressed, synthesising the required histidine for growth. Therefore any yeast colonies that grow on the 3'AT plates in theory have the two hybrid proteins interacting and this can be further tested by carrying out a β -galactosidase assay. Only those colonies arising from 'real' interactions, as opposed to *HIS3* revertants, will turn blue as they will express the *lacZ* gene.

This second assay therefore removes any potential false positives. The beauty of the yeast two hybrid system, when it is used to identify new binding partners, is that once such a real interaction is found the gene of interest can quickly be cloned and sequenced, as the cDNA fused to the Gal4 activation domain in the prey vector can easily be recovered.

The yeast two hybrid system has been used successfully to find protein partners (Bender et al., 1996; Durfee et al., 1993; Lesage et al., 1994), and even used to build networks of protein-protein connections through exhaustive searches (Fromont-Racine et al., 1997). It has also been used before to identify proteins that bind to WT1 (Johnstone et al., 1996; Johnstone et al., 1998; Wang et al., 1996). It is therefore hoped that the system can be used to identify other WT1-interacting proteins.

3.2 Testing the activity of WT1, as well as control proteins, in the Y2H system

Before a library screen can be carried out, the activity of WT1 in the Y2H system has to be tested. The reason for this is that some bait proteins can activate the expression of the reporter genes of the system on their own accord, without the need to bind to a prey. This could happen if the protein of interest, which is fused to the Gal4 DNA-binding domain, has an activation domain which could cause initiation of transcription. It was especially important to carry out this test on WT1 because WT1 has been shown to be able to activate transcription *in vitro* (see section 1.6.1). If WT1 does activate transcription of the system on its own, then it will not be able to be used as a bait to screen a library and instead it would have to be fused to the Gal4 activation domain and re-tested in the hope that it would no longer activate transcription on its own.

Whenever a transformation is carried out to see if there is an interaction occurring, there have to be controls added to ensure the system is working correctly. These controls have to be designed such that there is a positive and negative interaction so that growth on the plates can be checked accordingly. The controls to be tested as a positive interaction were WT1 +/+ used as both the bait and the prey, as it was known that this protein dimerises (Englert et al., 1995b). Another positive control to be tested in the system was the interaction between SNF1 (sucrose non fermenting gene) and SNF4, as these two proteins are known to interact (Celenza et al., 1989). The negative interaction to be tested was set up as WT1+/+ as the bait and SNF4 as the prey.

Bait	Prey	Number of colonies on plates	
		<i>AH</i>	<i>50mM AT</i>
WT1 +/+	None	0	0
WT1 +/+	SNF4	40	0
WT1 +/+	WT1 +/+	80	38
SNF1	SNF4	200	100
SNF1	WT1 +/+	200	0

Table 3.1: Testing the activity of WT1 and control proteins in the Y2H system. The control proteins produced the expected activities, and WT1 on its own does not activate the reporter genes, and so can be used as a bait for library screens to identify interacting proteins.

The activity of WT1 and the control proteins was tested in the yeast two hybrid system using the bait and prey vectors which had been constructed by Dr. Davies. The results produced are summarised in table 3.1. From the experiment carried out, it is shown that WT1 on its own cannot activate transcription of the reporter genes. When WT1 is used as the bait, but no prey vector is added, the yeast cannot even grow on the AH plates and this is because they lack leucine, which is normally provided by the prey vector. This therefore shows that even on AH plates there are no background colonies. The WT1 bait interacting with SNF4 prey shows that if the prey vector is now added the yeast can grow on AH plates as they now express leucine, but cannot grow on the 50mM AT plates as WT1 and SNF4 do not interact. This interaction not only acts as a negative control, but shows that WT1 cannot activate the reporter genes if there is no interacting prey protein present. The interaction between WT1+/+ used as both the bait and the prey can be used as a positive control as it is seen that yeast colonies transformed with these two proteins do grow on 50mM AT plates because the dimerisation of WT1 causes the two Gal4 domains to come together and express HIS3. The SNF1 and SNF4 proteins which interact with each other, also show a positive interaction in the system. From this test experiment, it can be seen that the system can be used to screen a cDNA library using WT1 +/+ as the bait.

3.3 Screening an M15 library

To try to identify proteins that interact with WT1 a cDNA library was to be screened. A flow diagram is illustrated to show the steps taken when screening a library (see figure 3.2) In preference to using a commercial kidney library, where WT1 only represents a small subset of the proteins present (see WT1 expression in the kidney, section 1.7.1), an M15 cDNA library was to be constructed and used. The M15 cell line expresses endogenous WT1 and it was thus hoped that by using this cell line, WT1 - and the proteins it interacts with - would be more highly represented in such a cDNA library compared to a whole kidney library, thus increasing the chances of isolating them.

3.3.1 Generating an M15 cDNA library

An M15 cDNA library fused to the prey expression vector is not commercially available and therefore it had to be made, with the help of Dr. Ladomery. PolyA mRNA was isolated from M15 cells and cDNA was made using oligo (dT) primers. The cDNA was then cloned into the prey expression vector using the HybriZAP two-hybrid cDNA gigapack cloning kit from Stratagene, as directed by the manufacturers. The cDNA library had an average insert size of 400bp, with the largest being about 1kb.

3.3.2 Identifying WT1-interacting clones

The M15 cDNA library was screened in the yeast two hybrid system to see if any proteins encoded by it interacted with WT1+/+. 270,000 colonies were transformed, as calculated by plating out a small proportion of the library onto AH plates, and out of these 20 colonies grew on 50mM AT plates. However the colonies that grew were very small and did not have the appearance of a convincing positive clone, as compared to the positive controls used. Nonetheless, the colonies were re-streaked onto fresh 50mM AT plates and a β -galactosidase yeast filter lift assay was carried out, to ensure that the clones were 'real'. Only 4 colonies, out of the 20, turned blue, and thus this confirmed the initial suspicion that the small colonies that had grown were false positives.

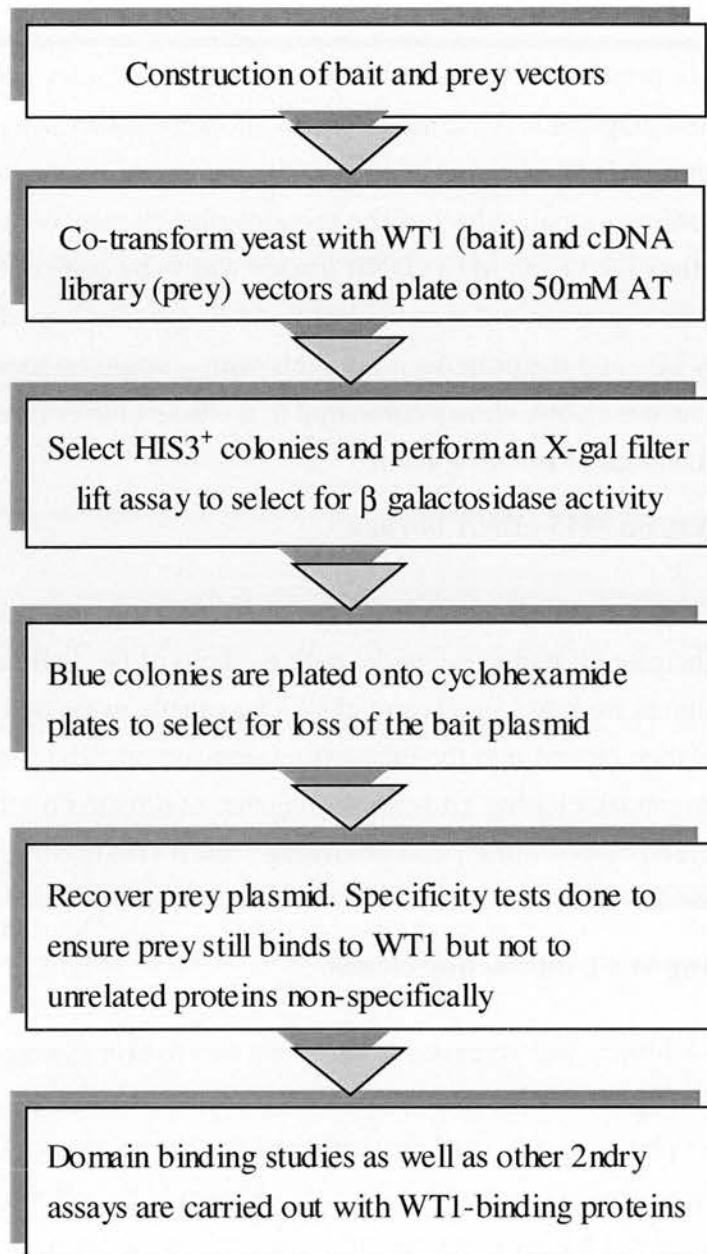


Figure 3.2: Flow chart showing the steps taken when using the yeast two hybrid system to screen a library for WT1-interacting proteins.

As the screen produced such few colonies, the plates were incubated for a few more days, in the hope that more colonies would grow, perhaps due to a weaker interaction. Another 60 colonies did grow, and out of these, 15 of them turned blue when assayed for lacZ expression.

3.3.3 Ensuring that the prey protein on its own does not activate the Y2H system

The next stage was to isolate the prey proteins which at this stage seem to bind to WT1, and check that they cannot activate the Y2H system on their own accord, without having to bind to WT1. To isolate the prey protein, the positive colonies are first streaked onto AHT plates (AH plates containing added tryptophan) and then plated onto cyclohexamide plates as described in section 2.2.7. By plating the colonies onto these plates, the yeast should lose the bait plasmid, because now that the yeast have tryptophan in the medium, they no longer require the bait plasmid to provide them with that amino acid. In addition the bait plasmid contains a cyclohexamide sensitive gene, which makes the originally cyclohexamide-resistant yeast sensitive to the drug when they are plated onto cyclohexamide plates. For these reasons the yeast lose the bait vector, retaining only the prey vector, which provides them with the required leucine. A β -galactosidase yeast filter lift assay can then be carried out on these colonies, and since these colonies no longer have the bait protein, they should not turn blue as they cannot express lacZ. If the colony does turn blue, it suggests that the prey protein can activate the system without having to bind to the bait, and is therefore a false positive. One of these colonies tested did turn blue and was therefore discarded as being a false positive. The isolation of the prey plasmid from the remaining 18 colonies was achieved as described in section 2.2.7. A digest of the prey clones isolated was carried out to determine the size of the cDNA cloned. Unfortunately out of the 18 potential clones, 8 seemed not to have an insert. This was rather surprising as it is known that WT1 must have interacted with a prey protein to be able to grow on 50mM AT plates.

3.3.4 Testing the specificity for WT1 interaction

The 10 remaining clones which did contain inserts were still candidates for binding to WT1 and so had to be re-tested in the yeast two hybrid screen against WT1 to ensure that the interaction was real, and against a negative control, such as SNF1, to ensure that the prey protein did not bind weakly to other unrelated proteins. This last control was added because it had been noted that some proteins promote transcription in concert with a number of different hybrid partners. This could be because the protein may have a surface that has low affinities for many different proteins, such as a large hydrophobic surface (Allen et al., 1995). Thus the potential WT1-binding proteins were tested to see if they bound to another unrelated protein, to ensure that this was not the case.

Unfortunately out of the 10 clones that were re-tested in the Y2H, none seemed to interact with WT1 again, and must have been false positives. Thus although the selection procedure is a dual one, so as to cut back on the number of false positives, several clones which were false positives did come through the screen. However from the start of the screen the colonies had looked doubtful as they were very weedy when they initially grew on the plates. Nevertheless some of the clones were sequenced to see if they encoded any known protein which could be of interest, and they could then be tested for their interaction with WT1 in an alternative way. However, the clones only matched to ESTs thus giving no clue as to what their possible role could be.

3.3.5 Summary of M15 screen

The screening of the M15 cDNA library did not isolate any WT1-binding proteins. The screening had not been carried out to saturation and if this had been done, then perhaps WT1-interacting clones could have been isolated. However it was decided that the library was not to be screened further as it did not appear to be an adequate library because the largest insert isolated was only 1kb, the average size being 400 bases. It is likely therefore that only small domains of proteins are present in the prey and thus the regions of the proteins required for WT1 binding may not be present. Also because the cDNA had been made using only oligo (dT) as primers, it is possible that many of the clones would only have contained the 3'UTR, again not providing the domain required to bind to WT1. It was decided that a commercially available cDNA library was to be screened instead.

3.4 Screening a commercial kidney cDNA library

3.4.1 Choice of library

The commercial library chosen to search for WT1-interacting proteins was a Clontech human foetal kidney MATCHMAKER cDNA library. This particular library was chosen because the mRNA source was from normal, whole kidneys of aborted Caucasian foetus' at 19-23 weeks of gestation. The age of the kidneys used to make the library coincides with the temporal pattern of expression seen for WT1 (Pritchard Jones et al., 1990) and it was thus hoped that in this library there would be high expression of WT1-interacting proteins which could be picked out in the screen. The library was made by oligo (dT) and by random priming so that as many cDNAs could be represented and so that the clones extended beyond the 3' end. The cDNA was fused to the pGAD10 vector containing the Gal4 transcriptional activation domain.

3.4.2 Identifying WT1-interacting clones

The library was screened with WT1 as the bait, and control proteins were also used as was done for the M15 library screen. The bait and prey vectors were transformed using the lithium acetate method and the yeast were plated onto AH and 50mM AT plates. This library screen produced 31 colonies that grew on the 50mM AT plates, out of 1.2 million clones screened. These positive clones were all picked and streaked out onto a fresh 50mM AT plate and left to grow at 30°C before a β -galactosidase yeast filter lift assay was carried out. As is shown in figure 3.3, the filter lift assay indicated that 21 out of the 31 colonies turned blue. It was therefore assumed that the clones that did not turn blue were false positive, and this may sometimes occur if there is a mutation occurring in the yeast which overcomes the nutritional problem. The clones that did turn blue were therefore potentially real interactions with WT1 +/+ and so have to be tested further to confirm this.

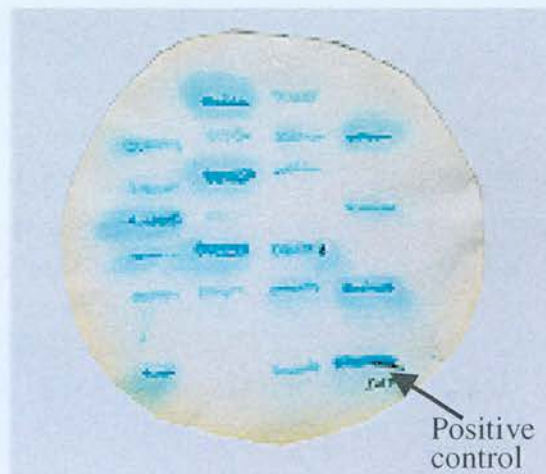


Figure 3.3: β galactosidase yeast filter lift assay. The 31 clones from the Clontech screen that grew on the plates that lacked histidine were streaked onto fresh 50mM AT plates. This β galactosidase assay yeast filter lift shows that 21 of these were potential WT1 interacting proteins.

3.4.3 Ensuring the clones are specific for WT1 interaction

At this stage it was very encouraging to obtain such a large number of clones which seemed to interact with WT1. The next stage was therefore to recover the prey plasmid so that the interaction could be re-tested in the yeast two hybrid screen against WT1 to ensure that the interaction was real, and against a negative control,

to ensure that the prey protein did not bind weakly to other unrelated proteins, as is described in section 3.3.4. The prey plasmids were recovered as described in section 2.2.7 and unfortunately when a digest was made from these plasmids, it was seen that 8 of the clones appeared to have no insert. The remaining clones had inserts varying from 700bp to 2.4kb. None of the clones turned blue when a β -galactosidase yeast filter lift assay was carried out after the colonies had been streaked onto AHT and cyclohexamide plates, showing that none of the colonies could activate the reporter genes on their own without having to bind to WT1.

The clones with inserts were re-tested in the two hybrid screen with WT1 and with SNF1. From the first batch of 5 clones tested one clone, clone E, showed no interaction with WT1 as shown in figure 3.4. Again, as for the M15 screen, this was a surprise finding as the clone was initially isolated due to its ability to interact with WT1. The other 4 clones tested did interact with WT1 but this time they all interacted with SNF1 too, as shown for clone A.

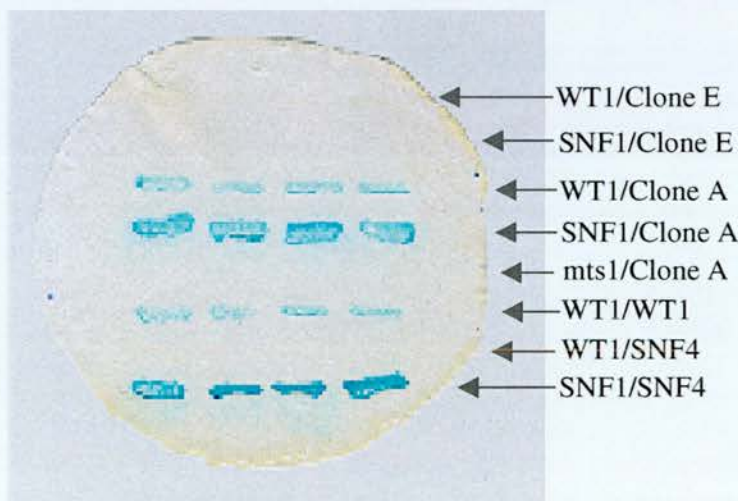


Figure 3.4: β galactosidase assay for the specificity test.

Clone E is an example of a clone that could not bind to WT1 when re-tested, despite having originally been isolated due to its interaction with WT1. Clone A is an example of a clone that could interact with WT1 when the specificity test was done, but it could also bind to SNF1. It could not, however, bind to mts1. The controls were added to ensure the assay was working correctly.

This was also disconcerting as it was not expected that so many proteins would bind non-specifically. To try to address this finding, the protein sequence of SNF1 was compared to that of WT1 to ensure that there was no common motif between the two proteins which may explain the reason for both binding to the same proteins. There apparently was no common motif between the two proteins, so it may be that either the 3-dimensional structure produced a similar fold in both proteins which would form a binding motif, or the clones were binding to both proteins non-specifically.

The validity of SNF1 acting as a negative control was now under doubt. Therefore to double check that these clones, as well as the clones not yet tested, were really not specific for WT1 it was decided that from then on they would all be tested against a range of different proteins and not just SNF1. These proteins were p53, msx1 and mts1, and these were chosen as they were already cloned into the bait vector and were readily available in the laboratory. From the 13 remaining clones - which were the ones that had grown on the plates lacking histidine, turned blue during the filter lift assay, and had an insert in the vector isolated - 8 no longer interacted with WT1 when re-tested and the other 5 did. Only one clone bound to WT1 and none of the other control proteins, and the 4 other clones all interacted with WT1 and SNF1 again, but not with the other 3 control proteins.

These results show that there is at least one clone, clone D, which is probably a real WT1-interacting protein as it specifically bound to WT1 but not any of the other control proteins. The other 4 clones (clones A, B, C and F) which did bind to WT1 and SNF1 but not to msx1, mts1 or p53 may be real interacting proteins but further tests would have to be carried out to see if the binding to SNF1 was non-specific.

3.4.4 Domain binding studies with the possible WT1-specific clones

Domain binding studies

Once a protein-protein interaction has been found, the yeast two hybrid system can also be used for domain binding studies, which maps the site of interaction. These studies test the binding of the candidate prey protein, isolated in a screen, to new bait proteins which now only contain domains of WT1 instead of the full length protein. Those domains of WT1 which are tested and continue to show a positive interaction with the candidate prey, are therefore involved in the binding interaction.

Domain binding studies were carried out on clones A and C, as random representatives of the four clones A, B, C, and F, to be able to determine whether these clones were binding to WT1 non-specifically or whether they were binding to both WT1 and SNF1 due to the presence of a common 3 dimensional motif. Thus if

clone A and C bound to the same domain of WT1, it would suggest that there was a 3 dimensional motif on WT1 that was being recognised by clones A, B, C and F which was also present on SNF1 thus accounting for these clones not being as specific as clone D when binding to WT1. Clone D was also used in these domain binding studies to see whether all 4 isoforms of WT1 could bind to it, and to determine which domain of WT1 was responsible for binding.

Cloning the bait plasmids for domain studies

The domain studies are carried out in the yeast two hybrid screen using deletion constructs of WT1 as well as all four isoforms of WT1 as baits. The cloning of all these baits was carried out by Dr. Davies. The following are the deletions constructs of WT1 that were used, and they are also depicted on figure 3.5a :

Nco5': this is a deletion construct containing the first 182 amino acids of WT1 (or exon 1 and most of exon 2). This was cloned by digesting the WT1+/+ bait with NcoI. The smaller of the 2 fragments was purified and ligated into pAS1-CYH2 bait vector, which had been digested with NcoI.

Nco3'+KTS : This clone contains the remaining segment of WT1+/+ not covered by *Nco5'*. The larger of the two fragments produced when WT1+/+ was digested with NcoI was cloned into the bait vector as for *Nco5'*.

Nco3'-KTS : This deletion construct is the same as *Nco3'+KTS* but it has the -KTS splice variant, and so was cloned in the same way too but originating from WT1+/-.

Nard3: This deletion construct consists of WT1+/+ that has had amino acids 38-159 removed. It is constructed by carrying out a NarI digest on WT1+/+ and re-ligating the large fragment into the bait vector.

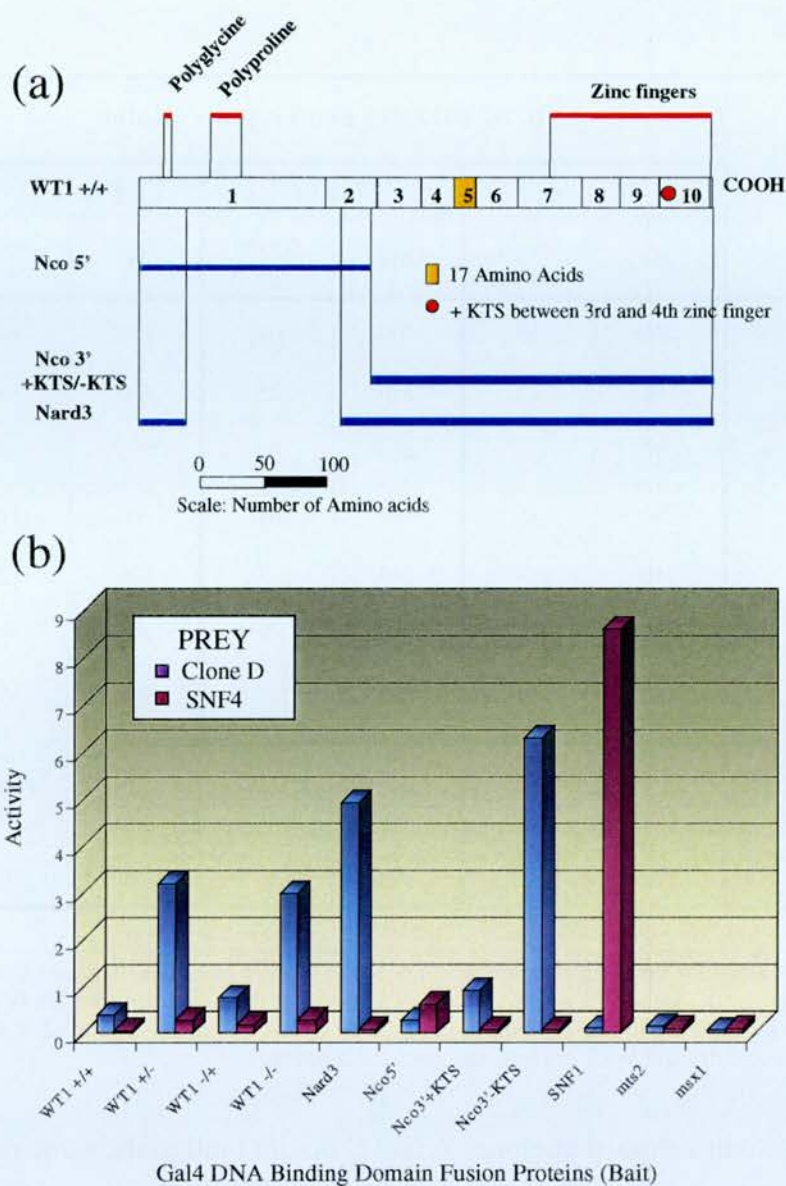


Figure 3.5: WT1 domain binding studies with Clone D. Panel (a) is a schematic showing the WT1 deletion constructs used in the assay. Panel (b) is a graph showing the binding of Clone D, and the control SNF4, to the different WT1 isoforms and deletion constructs as well as to non-related proteins.

Domain binding studies using the yeast two hybrid system

The clones A, C and D were tested for this domain binding study, by co-transforming each clone with all of the WT1 isoforms or deletion constructs in turn. The bait plasmids that continued to bind to the candidate clones would contain the region of WT1 required for interaction. The results are summarised in table 3.2.

Bait	No. of colonies growing on plates					
	Clone A		Clone C		Clone D	
	AH	50mM AT	AH	50mM AT	AH	50mM AT
WT1 +/+	200	10	400	100	750	67
WT1 +/-	100	12	210	25	500	350
WT1 -/+	200	30	400	50	783	250
WT1 -/-	280	16	400	100	700	420
Nard3	140	14	240	5	145	128
Nco3' -KTS	350	50	600	0	738	500
Nco3' +KTS	110	14	200	0	930	350
Nco5'	210	0	240	58	1000	0
p53	400	0	600	0	400	0
mts1	320	0	400	0	270	0

Table 3.2: WT1 domain binding studies. Clones A and C were tested in these studies to see whether they bind to the same WT1 domain, to assess whether their interaction with WT1 is specific. Clone D was used in this domain binding study to see if it could bind to all 4 WT1 isoforms, and to identify the WT1 domain required for binding.

These binding studies show that clones A and C do not bind to the same region of WT1, as clone A binds to the C-terminal domain of WT1 while clone C binds to the N-terminus of WT1. The difference in the number of colonies growing for clone A compared to clone C suggests that clone A may be slightly toxic to the yeast. These results suggest that the interaction between clones A, and C, with WT1 are likely to be non-specific as they do not bind to the same 3D motif which would otherwise account for the binding of these clones to SNF1.

The domain binding studies carried out on clone D show that it can indeed bind to all 4 isoforms of WT1 and it seems to not be able to bind Nco5', thus indicating that exons 1 and 2 are not involved in the interaction. To give a better idea of the strength of interactions between clone D and the different isoforms as well as the deletion constructs, a liquid β -galactosidase assay was carried out as described in section 2.2.6 (Estojak et al., 1995).

The graph shown in figure 3.5b shows the results obtained when the liquid β -galactosidase assay was carried out. The data in the graph represents an average of 2 independent experiments, each itself an average of the results obtained from 3 separate colonies. The graph shows that clone D does bind to all four isoforms above background levels, where background is taken as the equivalent of the bait protein binding to SNF4. However, the graph clearly shows that the interactions between clone D and the -KTS isoforms of WT1 are a lot stronger than the interactions with the +KTS isoforms.

The deletion constructs then show that exons 1 and 2 are not required for binding as Nco5' does not show an interaction with clone D. The deletion construct Nco3'-KTS showed the strongest interaction with clone D, and when compared to the binding with Nco3'+KTS, it shows that the +KTS results in a reduced affinity for the interaction. The reason why the interaction between clone D and the Nco3'-KTS is stronger than for the wild type -KTS isoforms may be due to the fact that the deletion construct removes the repression domain of WT1. This domain may interfere slightly with the transcription of the reporter genes by Gal4. The removal of the repressor domain in the Nco3'+KTS may also explain why this deletion construct has a slightly higher binding affinity to clone D compared to the wild type WT1 +KTS isoforms. However the reason why Nard3, which also has +KTS, has a higher affinity for clone D compared to the other +KTS clones is not yet understood. It would be of interest to see if the -KTS version of Nard3 would have an even higher affinity. These results suggest that clone D binds to the C-terminus of WT1, maybe via the leucine zipper domain, but more likely to the zinc finger region because the insertion of +KTS seems to reduce binding.

3.4.5 Sequencing the clones

Although at this stage it seemed that clones A, B and C were non specific for WT1 binding, they were still sequenced together with clone D. This was done in case there was any of these clones encoded a protein which may be of interest. Clone A, which was 2.4kb, appeared to match fibrillin, but the cDNA clone was out of frame so that instead of encoding for fibrillin, a non-sense protein was made. This non-sense protein may now possibly fold in such a way that it could interact with WT1 but because it is likely to be non-specific binding, this clone was now abandoned as a possible WT1-interacting clone.

Clone B, whose sequence matched a human EST, AA 083180, from a Stratagene endothelial cell line and clone C which had no match in the database were also

discarded as possible WT1-interacting proteins because their binding to WT1 seemed to be non-specific and their sequence did not highlight any motifs which could be of interest.

Clone D, which was the only specific WT1-binding clone and was 1.2kb in length, matched the human EST D14661. This EST had no motifs that appeared in the database and so it yielded no clues as to its possible function, nor why it was binding to WT1. However, because this protein did seem to bind to WT1 in the yeast two hybrid system, it could still produce some useful information if characterised further. Further test therefore had to be carried out on clone D which from now on was to be called NAL (novel associating ligand), as it was a novel WT1-binding protein.

3.5 NAL, a candidate WT1 binding protein

As NAL was now a candidate WT1 protein, other tests using the yeast two hybrid screen were carried out to provide more evidence about the specificity of its interaction with WT1.

3.5.1 NAL can bind to WT1 when expressed as the bait or the prey protein

NAL was isolated from the Clontech library screen when WT1+/+ was used as the bait. As a indication of whether the association between the two proteins is 'real', a domain swap experiment can be done. In these experiments, the selection procedure is swapped around such that the cDNA encoding for NAL now becomes fused to the DNA-binding domain of Gal4, instead of to the activation domain. Now NAL acted as the bait, and WT1+/+ was the prey. The WT1-/- isoform was also tested to see if this too could bind to the NAL when it acted as a bait.

Table 3.3 shows the results of the yeast two hybrid screen carried out with NAL as the bait and WT1 as the prey, together with controls. This experiment shows that the interaction between NAL and WT1 is very likely to be real as the association between the two proteins can be found whether NAL is acting as the bait or the prey. The failure to see an association between NAL, as the bait, and WT1-/-, as the prey, does not mean that the interaction is not real with this isoform, as when WT1-/- is used as the bait then there is an association seen. The reason behind not seeing an association when the prey and the bait have been swapped could be because the conformation of the two proteins in their new environment may have altered slightly such that the binding region of the protein becomes masked. This experiment can therefore only have a positive interpretation in that if an interaction occurs then it

provides more support for the association. The Y2H data produced for NAL, points towards the interaction between WT1 being real and specific.

Bait	Prey	No. of colonies on plate	
		<i>AH</i>	<i>50mMAT</i>
NAL	WT1+/+	800	120
NAL	WT1-/-	1200	0
NAL	SNF4	325	0
WT1+/+	NAL	1200	400
WT1-/-	NAL	1000	400
SNF1	NAL	1000	0
SNF1	WT1+/+	900	0
SNF1	WT1-/-	1200	0
SNF1	SNF4	275	130
WT1+/+	SNF4	250	0
WT1-/-	SNF4	300	0

Table 3.3: NAL can bind to WT1+/+ when acting both as the prey and the bait.

3.5.2 NAL binds to the -KTS, zinc finger region of WT1

Section 3.4.4 shows that NAL bound to WT1 with a higher affinity for the -KTS isoforms, and it appeared that NAL was binding to the zinc fingers and not to the leucine zipper motif because the +KTS disrupted binding. A quick Y2H screen with the zinc fingers, both +KTS and -KTS, of WT1 was carried out, to see if NAL could bind to them without the rest of the WT1 protein being present. The results obtained are in table 3.4.

Bait	Prey	No. of colonies on plate	
		<i>AH</i>	<i>50mM AT</i>
ZF +KTS	NAL	1200	0
ZF -KTS	NAL	800	700
WT1+/+	NAL	1000	50
SNF1	NAL	1000	0
ZF +KTS	SNF4	1000	0
ZF -KTS	SNF4	500	0
WT1+/+	WT1+/+	1600	300
WT1+/+	SNF4	1200	0
SNF1	SNF4	1600	800

Table 3.4: NAL can bind to the zinc fingers (ZF) of WT1 but only if they are -KTS.

This experiment indicates that NAL can bind to just the zinc fingers of WT1 so long as it lacks the 3 amino acids, KTS. The insertion of the +KTS must alter the conformation of the zinc fingers such that NAL can no longer bind to them. However the fact that NAL can bind to WT1 +KTS isoforms, although weaker than the -KTS isoforms, suggests that the remainder of the WT1 protein stabilises the interaction between NAL and the +KTS zinc finger motif, thus allowing for some binding. These results suggest that NAL preferentially binds to the -KTS zinc fingers of WT1. By binding to the zinc fingers of WT1, NAL may be interfering with the DNA and/or RNA properties of WT1.

3.6 Conclusions

The aim of using the yeast two hybrid system to screen cDNA libraries was to identify WT1-interacting proteins. If known proteins were identified, then through virtue of knowing their function, the consequences of them interacting with WT1 could be addressed and this might increase the understanding of the function of WT1 at the molecular level.

The first library screen performed on the M15 cDNA library constructed, did not identify any WT1-interacting clones. This was unfortunate, but may have been due to the library containing very short cDNAs, which because the library was made from oligo (dT) primers, might not have included much more than the 3'UTR of

clones. The Clontech library screen did isolate some WT1-binding proteins. In the end it seems that out of 5 potential WT1-binding proteins, only one clone ended up being specific for WT1, as the rest all bound to an unrelated negative control, SNF1. This problem with specificity has been reported before (Fromont-Racine et al., 1997) and they suggest that perhaps it is due to the presence of the HA epitope tag which is present in the bait protein. The one protein that did bind to WT1 specifically was NAL, which is a novel protein. It was a shame that no known proteins were isolated in the screens, especially proteins which have already been documented to bind to WT1, such as WT1 itself. There are several possible reasons as to why no known proteins were isolated, and indeed why such few clones were isolated from the screens, as outlined below.

One reason, as already mentioned for the M15 library, was that the cDNAs in the library are not full length. Although the Clontech library used both oligo (dT) as well as random priming to synthesise the cDNA, and thus have more than just the 3'UTR, the clones may still be lacking the domains needed for binding to WT1. For example, for WT1 to dimerise the N-terminus is required, and if this motif is not present in the cDNA made from the WT1 transcript, then the Y2H screen will not isolate the WT1 transcript in the library.

A problem encountered when carrying out the screen was that clones that originally had been isolated because they interacted with WT1, no longer did so when the clone was isolated and re-tested. One possibility for this occurrence is that when the yeast are transformed, some may pick up more than the one prey plasmid. It is likely that out of the transformed prey plasmids, only one will bind to WT1 and thus when the vector is isolated, it is possible that the other prey present in the yeast is captured. Therefore when this clone is re-tested it will no longer be able to activate transcription as it was not the original prey that was binding to WT1. This hypothesis can also explain why clones that seemed to activate the system had empty vectors i.e. the yeast had been transformed with two preys, one of which had an WT1 binding insert - thus causing transcription of the reporter genes - and the other one which was empty, but happened to be isolated in preference to the WT1-binding clone. This hypothesis could have been tested by carrying out mating type assays, where the prey plasmid does not have to be isolated as the colony is directly streaked to a fresh plate and mated with a strain of the opposite mating type. In these mating type assays, the clones should still be able to interact with WT1.

Another problem with the system was that the screen may not have reached saturation and this was certainly true for the M15 library screen where only a quarter

of a million clones were screened. Over 2 million clones had been screened in total for the Clontech library. Although in theory this would have screened the 2 million clones present in the library, it is highly unlikely that all clones would have been screened. It is also possible that when the library was being amplified, some clones may have been selected for and therefore enriched, whilst others may have been selected against, especially those that could be harmful towards the bacteria. For this reason a lot of potential candidate WT1-interacting proteins may not have been represented in the libraries. Also some of the cDNAs which are represented in the library may be out of frame, as was seen for clone A, which was a fibrillin clone out of frame. Only one sixth of the cDNAs will be in the correct frame for translation, and so a lot of nonsense proteins will be present in the library.

Another possibility as to why so few clones were isolated is highlighted by section 3.5.1. In this experiment, it was found that NAL could bind to WT1 +/+ whether NAL was used as the prey or the bait. However, NAL could only bind to WT1 -/- if NAL was the prey but this did not mean that the interaction was weaker because the quantitative β -galactosidase assay shows that in fact the WT1 -/- isoform interacted better with NAL than the +/+ isoform. Thus the conformation of WT1 may be altered slightly depending on whether it is fused to the Gal4 DNA-binding domain or to the Gal4 activation domain, and this may be sufficient to prevent some interactions from occurring. Therefore if the screen had been performed with WT1 as the prey instead of the bait, then other proteins which can only bind to WT1 in this orientation may have been isolated. This has in fact been shown by Davies *et al.* as U2AF65, which has been identified as a WT1-interacting protein (Davies *et al.*, 1998), can only bind to WT1 if the latter is used as the prey.

All these reasons can explain why the screens did not identify other WT1-interacting proteins. It would be possible, in an attempt to identify other WT1-interacting proteins, to continue screening the libraries, to ensure that the screens were carried out to saturation, or to screen other available libraries. It would also be feasible to carry out the screens using WT1 as the prey, but this would mean having to reclone the libraries into the DNA binding domain vectors. Alternatively a different approach to identify WT1-interacting proteins could be carried out, such as screening lambda expression libraries. However, all these possibilities were not carried out, because it was decided that the Y2H screening had already identified a potential WT1-interacting protein - NAL. NAL is a good candidate because it binds to WT1 specifically, without binding to unrelated proteins and it binds to WT1 +/+ whether it is expressed as the bait or the prey, thus providing more support to suggest that the interaction is 'real'. NAL seems to be isoform dependent, binding

more strongly to the -KTS isoforms. The WT1 zinc fingers seem to be the motif NAL binds to, and this suggests that perhaps NAL may interfere with the DNA and/or RNA functions of WT1. Thus although NAL is a novel protein and so does not provide any immediate clues as to why it binds to WT1, it is still of interest and should be characterised further in the hope that it may eventually shed some light on WT1 function.

Chapter 4

Identification of a Novel Protein

Chapter 4: Identification of a novel protein

As detailed in chapter 3, NAL was identified as being a potential WT1-binding protein. It was of interest to characterise this protein further to gain as much information about it as possible, and so the clone was fully sequenced and mapped.

4.1 Sequencing NAL

The sequencing of NAL that had been described in chapter 3 had only covered the first 300 bases of the clone and so now it was sequenced fully to see if there were any motifs within the sequence which could point to a possible function. The clone was sequenced in the forward and reverse orientation to ensure the fidelity of the sequence produced. Each section of sequencing run was also carried out in duplicate. Figure 4.1 shows the sequence obtained together with the primers used (the primers are also detailed in the appendix).

4.2 Database matches for NAL

4.2.1 NAL partially matches EST D14661

The sequence obtained for NAL was run through the GenBank database to see if it matched any known sequences. The EST D14661 produced a 60% identity match. However, it was noted that for the first 445 bases of NAL, the match was 100% identical, after which there seemed to be a breakpoint where there was no longer any real homology - see figure 4.2. This was highly suggestive of one of the two sequences originating from a chimaeric clone, formed by the fusion of two non-related cDNAs. This possibility had to be clarified to determine which of the two was from a chimaeric clone.

4.2.2 Do either NAL or the EST originate from a chimaeric clone?

To address the question of whether NAL, or the EST, originated from a chimaeric clone, RT-PCR was carried out on RNA from M15, COS and HeLa cells (see appendix for primer sequences). For both the EST and NAL, the PCR was carried out across the breakpoint (primers M729 and M730 for the EST, and primers M729 and M731 for NAL) and across most of the whole length of the clone (primers M717 and M728 for the EST and M717 and M718 for NAL). A PCR was also carried out for the region that was perfectly matched between the two, called the N-terminus (primers M717 and M715), as a positive control as well as for the C-terminus of NAL (primers M716 and M718) to ensure that this sequence was also an expressed

sequence. Figure 4.3a illustrates which primers were used to amplify the different segments of the clones.



Figure 4.1: NAL sequence showing primers used to sequence it. Primers H792 and H793 originate in the pGAD10 vector. The NAL clone is not full length as it does not have a start or a stop codon.

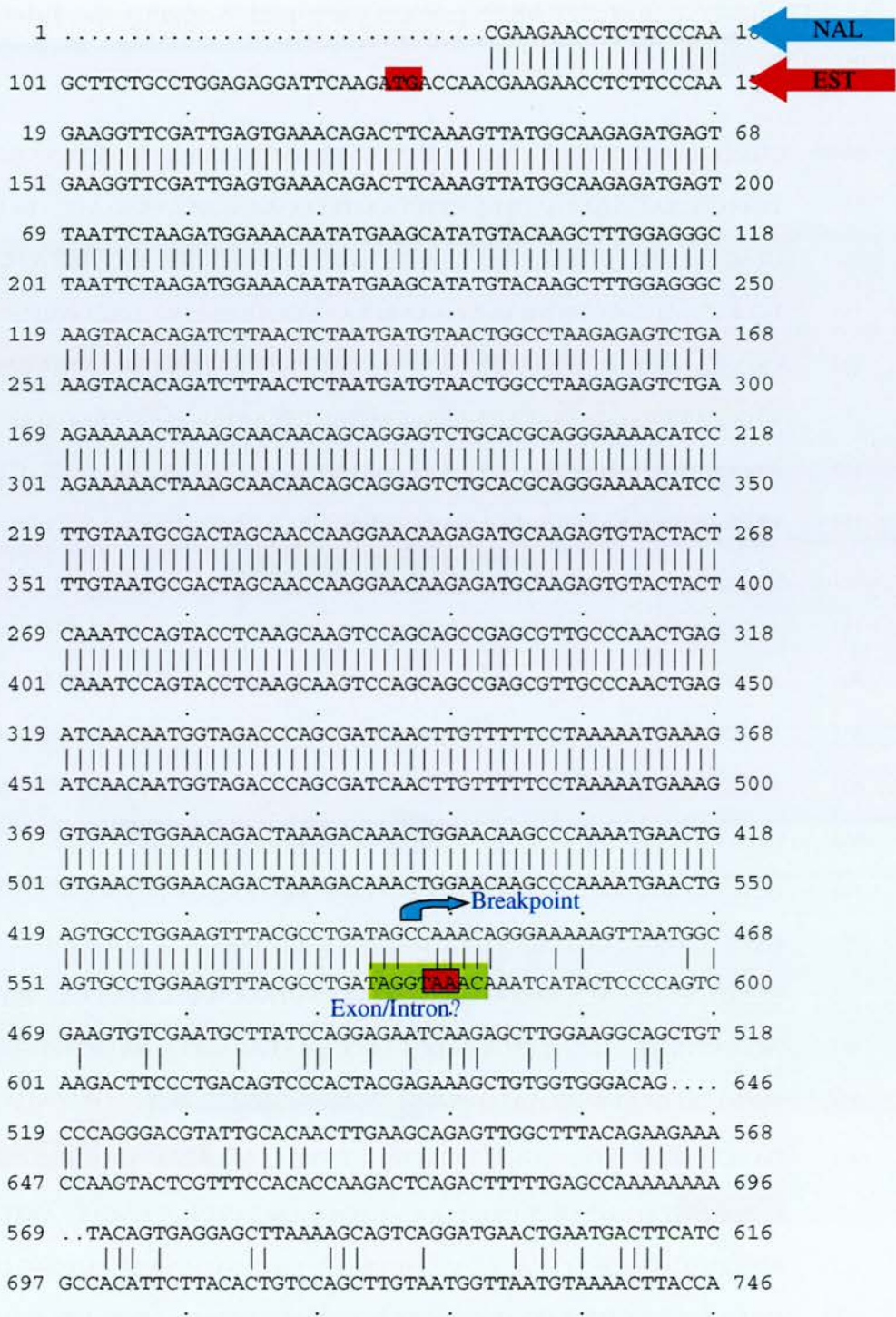


Figure 4.2: Sequence comparison between NAL and the EST, D14661.

It can be seen that the sequence is 100% identical until base 445 of NAL, after which there is a breakpoint where there is no real homology. At the breakpoint the EST has a 5' splice site consensus sequence, followed by a stop codon. The EST sequence continues on (not shown) and at the end it has a poly-adenylation site (AATAAA).

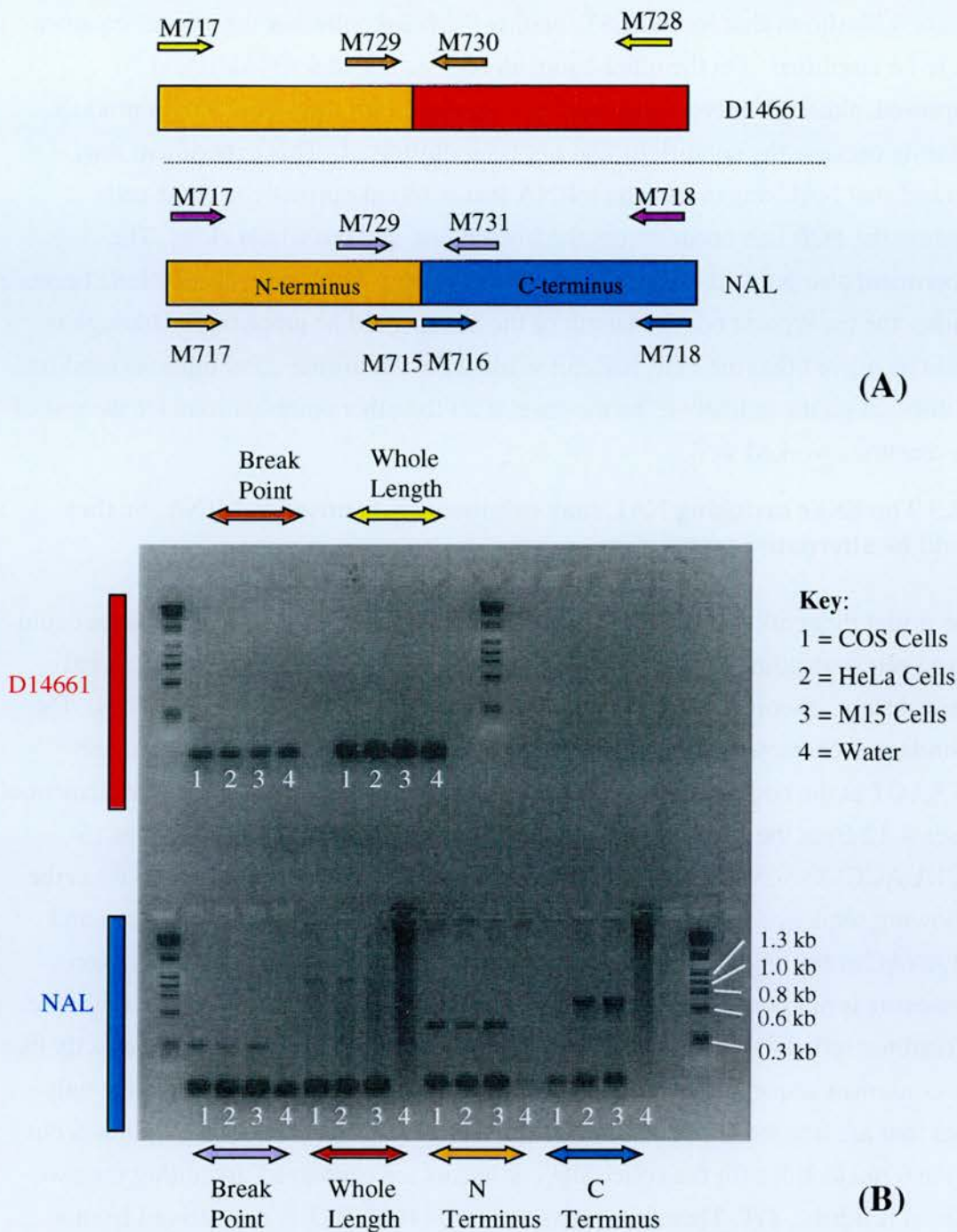


Figure 4.3: RT-PCR experiment to determine whether NAL and D14661 are transcribed in cells. Panel A shows the primers used to amplify the various segments of the EST and NAL clones. Panel B shows the products of the RT-PCR run on a 1% agarose gel. For NAL, products can be seen for every combination of primers, although the whole length product is very faint. For the EST on the other hand, no products can be amplified for any of the cell lines used.

Figure 4.3b shows that for the EST, neither the breakpoint nor the whole sequence could be amplified. On the other hand, all regions tested for NAL could be amplified, although only a faint band was produced for the whole length product, possibly because the conditions had not been optimised. This experiment thus showed that NAL was in fact the mRNA that is found normally in these cells because the PCR can occur across the breakpoint, and the whole clone. The experiment also pointed towards the EST originating from a chimaeric clone because neither the breakpoint nor the length of the clone could be amplified. Although it could be argued that the PCR had not worked for the primer combinations used for D14661, it seems unlikely to be the case as all the other combinations for the rest of the reactions worked well.

4.2.3 The ESTs matching NAL may originate from unspliced RNA, or they could be alternative transcripts

The initial thoughts that the EST, D14661, originated from a chimaeric clone could have been misleading as when the sequence across the breakpoint was analysed more closely, it seemed to resemble an exon-intron boundary (see figure 4.2). This boundary is characterised by the sequence (A/C)AG at the end of the exon and GTAAGT at the beginning of the intron. This sequence is the perfect complement of bases 4-12 from the 5' end of the U1 snRNA component of the spliceosome: 5' ACUUACCUG 3', which guides selection of the 5' cleavage site. The EST has the following sequence at the breakpoint: TAG at the end of the presumed exon and GTAAAC at the beginning of the proposed intron. The initial base of the exon consensus is not conserved as the G at position 12 of U1 RNA is thought not to be critical for splice site selection. Bases 5 and 6 of the EST intron are not exactly like the consensus sequence either, but some bases are not 100% conserved - the only ones that are, are the first two bases of the intron -GT (Jackson, 1991). Thus 6 out of the 9 nucleotides for the splice site consensus are conserved, including the two critical residues, -GT. Therefore perhaps this D14661 EST is not derived from a chimaeric clone after all but is instead an unspliced RNA. Further reason to believe that the EST has not arisen from a chimaeric clone is that when the EST D14661 is run through the database, it matches other EST clones which have exactly the same sequence, even after the breakpoint, and it is extremely unlikely that they all originate from chimaeric clones.

To see if the C-terminus of the NAL clone, which did not match the D14661 EST, was homologous to any other clones, it too was put through the database. Intriguingly it matched an EST (H90530), amongst others, but again there seemed to

be a breakpoint in H90530 where there is a sudden loss in homology after a good match. This time the initial sequence of H90530 doesn't match NAL, and then after a breakpoint does so perfectly as illustrated in figure 4.4. If this breakpoint is examined, it shows that it is an intron-exon boundary. This has the consensus (Y)₁₀NYAG for the sequence of the intron and a G for the start of the exon, (where Y= T or C). The breakpoint of where the clone H90530 begins to match NAL is CTCTTTCCTTTGCAG at the end of the intron and a C at the start of the exon. This therefore matches 100% for the end of the intron sequence. The start of the exon sequence does not match but again there is not 100% conservation for all bases, except for the AG at the very end of the intron (Jackson, 1991), which is what we have. Thus again there seems to be an unspliced RNA in the database. Unfortunately the 5' sequence of the H90530 intron does not extend far enough to see if it reads into the 3' end of the D14661 intron.

The picture that is emerging is therefore that NAL is a bone fide cDNA, which is made up of several exons - the first few being the same as the exons in D14661 and others being the same as the exons in H90530. The introns of the primary RNA transcript of NAL would therefore consist of the introns in D14661 together with those of H90530. The picture can however be more complicated than that. Although it is not uncommon for the database to have ESTs originating from unspliced RNA, it seems unlikely that so many clones in the database which are like D14661, but shorter, are unspliced. Therefore if the sequence of D14661 is looked at further, it shows that in fact it could encode an open reading frame (ORF) that goes beyond the exon-intron boundary until it reaches a stop codon which is found straight after the breakpoint. The rest of the sequence would then be 3'UTR, at the end of which there is a poly-adenylation site (AATAAA). It now appears that the 5' splice site of this primary RNA transcript could either be used, and so by removing the intron and joining on the next exon - followed by a 3'UTR - it would produce NAL, or it could be read through and so produce a smaller ORF, with its own 3'UTR, such as that just described for D14661 (see figure 4.5). This phenomenon of having alternative 3'UTR and polyA sites within a transcript has been described for other genes (Edwards-Gilbert et al., 1997; Lee and Maihle, 1998).

h90530.gb_est4 x NAL.seq ..

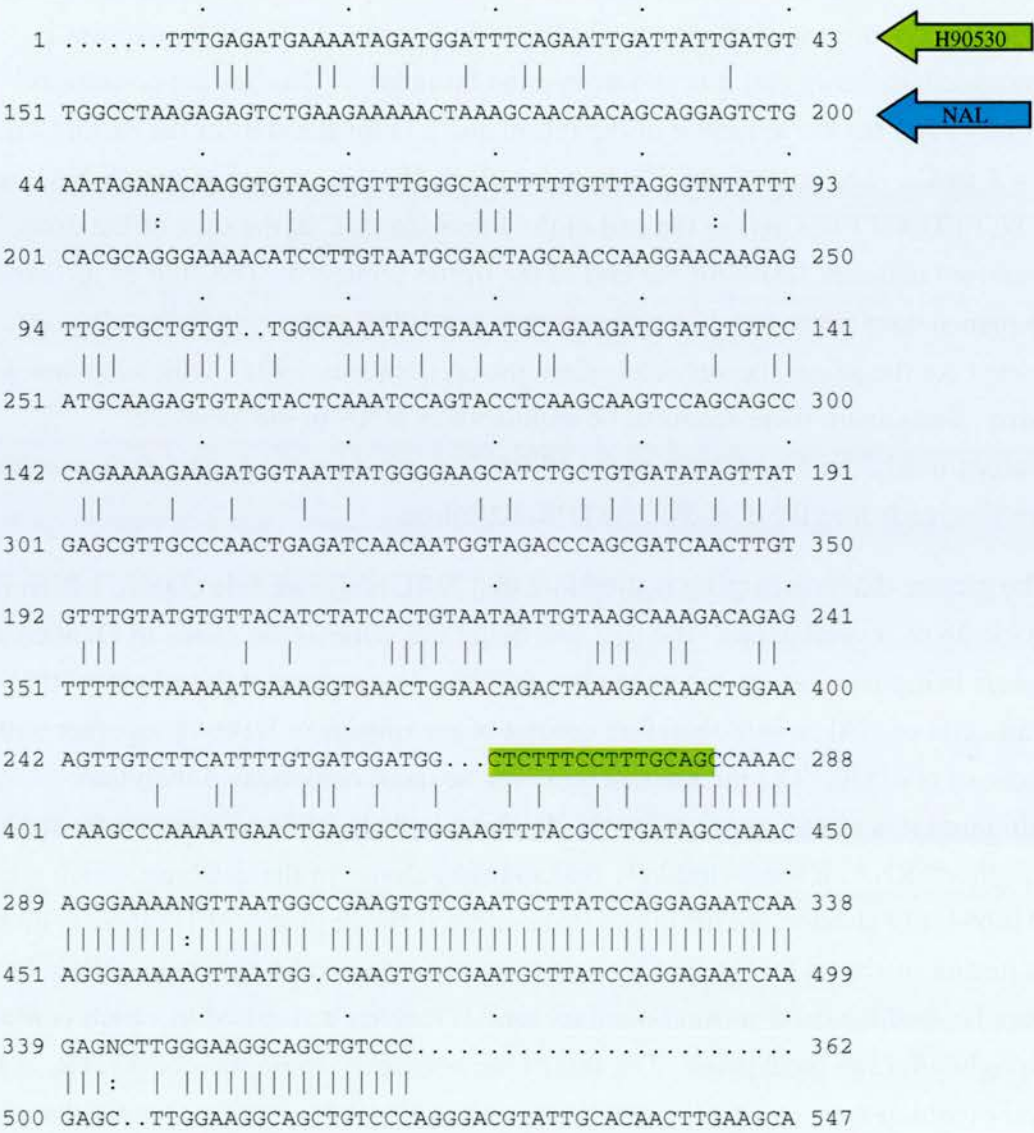


Figure 4.4: Sequence comparison between the EST, h90530, and NAL. The two sequences only match from base 445 of NAL. It is likely that the h90530 clone is unspliced, because as highlighted by the green box, there is an intron/exon consensus sequence.

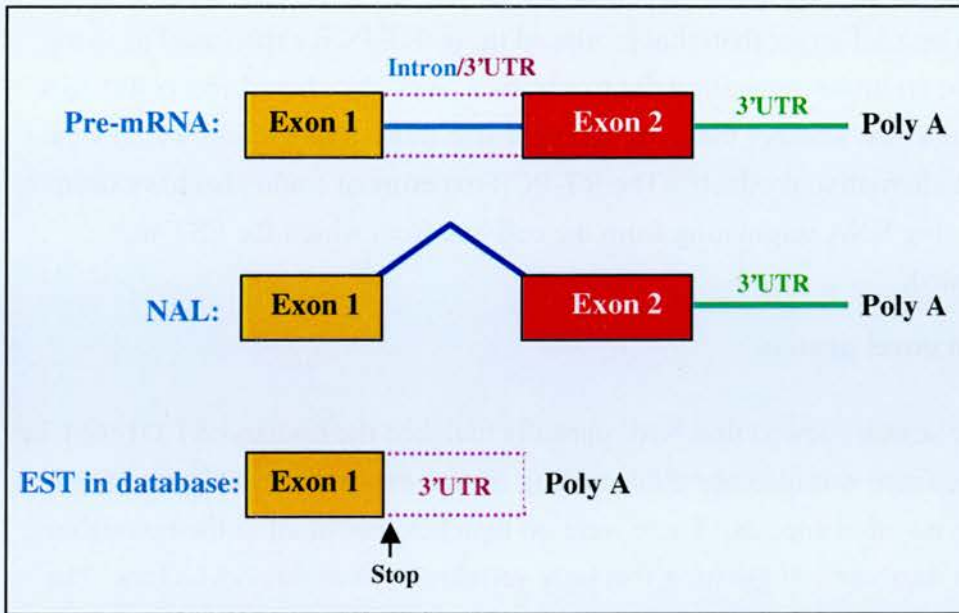


Figure 4.5: Schematic showing the model proposed for the formation of two alternative transcripts. From the pre-mRNA 2 alternative transcripts can be produced to form the EST and NAL. As the genomic sequence is not known, for simplicity's sake, the group of exons in NAL which match the EST will collectively be called exon 1, and the group of exons which do not match the EST will collectively be called exon 2. In the proposed model then, the intron could be spliced out, and exon 1 and 2 joined to form NAL, which has its 3'UTR and polyadenylation site. Alternatively the splice site at the junction between exon 1 and the intron could be read through, such that the intron becomes a terminal exon, which has its own 3'UTR and polyadenylation site which are different to those of NAL.

Initially it would appear that the RT-PCR experiment carried out would not back this hypothesis up, because if there were to be read through to produce D14661, then there should have been a product formed. As the 5' splice site consensus sequence is well conserved, it likely to be spliced in preference to being read through, but all this will depend on levels of poly-adenylation factors, splicing factors and termination factors within a particular cell. Therefore the transcripts may be cell type specific and as only three cell lines were tested for the RT-PCR experiment, it cannot be excluded that D14661 might not be amplified in other cell lines. Another possibility is that D14661 is present in the cell lines tested but the mRNA may not be very stable, and so cannot be amplified. Although not carried out, the best way to have tested this hypothesis would have been to carry out the PCR on genomic DNA. If D14661 does have this exon-intron boundary, through which splicing may or may not occur, then this sequence should be the equivalent to the genomic DNA, and so if the PCR is carried out across the breakpoint then a small product should be seen. For NAL, on the other hand, the breakpoint PCR product produced from genomic

DNA should be a lot larger than that produced in the RT-PCR experiment as there would now be an intron separating the two primer pairs. Northern blots could also be carried out to see whether there are different size transcripts present which would represent the alternative products. The RT-PCR experiment could also have been carried out using RNA originating from the cell line from which the EST was initially isolated.

4.3 NAL is a novel protein

The database search showed that NAL partially matched the human EST D14661 in the database. There was also one clone arising from a mouse and one from rat cell line but from no other species. There were no matches seen at all in the invertebrate nor the fungi databases, suggesting that only vertebrates have this NAL clone. The Swissprot database was searched to see if there were any homologous proteins but none were found. The motif database also showed that the only motifs known in NAL were those of phosphorylation sites; amidation and glycosylation sites but apart from these there were no motifs which could be associated with any particular function. The only recognisable motif present in NAL is that of glutamine rich stretches, which are often associated with transcription factors. These database searches conclude that NAL is a novel protein which has not been characterised before.

4.4 Isolating a mouse cDNA homologue of NAL

The mouse clone found in the database that matched the human EST, D14661, started at the position equivalent to base 672 in the human D14661 clone. This mouse clone therefore matches the EST after the breakpoint, and so is either part of the 3'UTR of the EST, or if the human EST did arise from a chimaeric clone then it is from the cDNA that does not match NAL. Either way the mouse clone in the database did not match the NAL transcript, and thus a library screen was to be carried out to isolate the mouse homologue of NAL.

4.4.1 Screening a mouse cDNA library

The library screen was to be performed on a mouse cDNA library to obtain the mouse homologue of NAL. Initially an RT-PCR was carried out on whole mouse mRNA, E=11.5d, to ensure that there was a mouse homologue of NAL present at this age, before carrying out the screen. The primers to be used were those that would amplify the breakpoint, the length of the whole clone as well as the N- and C-terminus of NAL, as these were the primers that were used for section 4.2 As

shown in figure 4.6 products were amplified for each of the primer pairs used, but again the whole length product was very faint due to the conditions used. This RT-PCR experiment suggests that there could be a mouse homologue of NAL in the E=11.5d library. This experiment also suggests that the mouse and human NAL sequences must be fairly similar as the human oligos used could also recognise the mouse sequence.

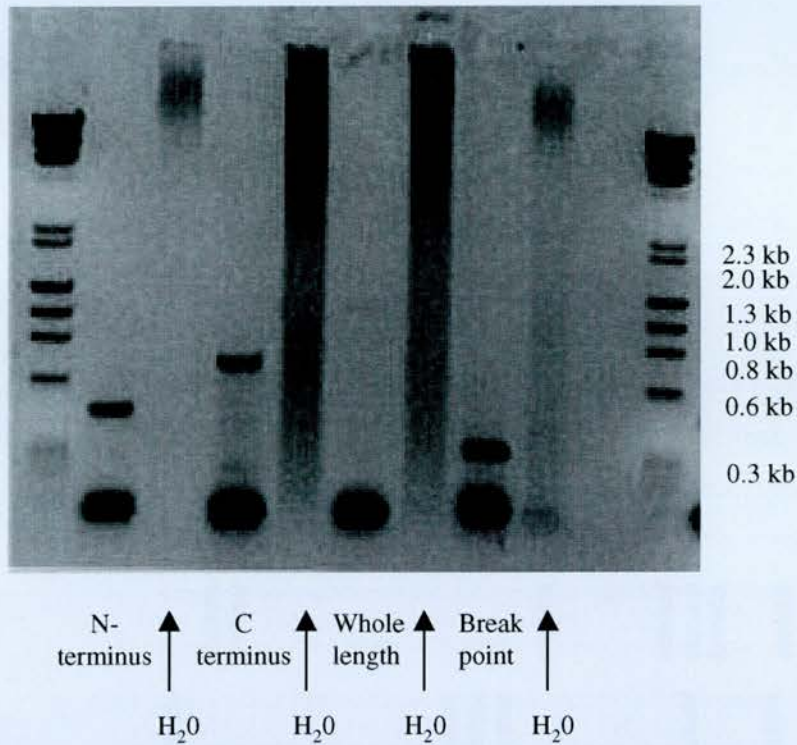


Figure 4.6: RT-PCR on whole mouse RNA (E=11.5d).

Primers against the N-terminus; C-terminus; whole length; and across the breakpoint of NAL were used, as depicted in figure 4.3a. The products of each amplification can be seen, except for the whole length where only a faint band was produced due to the PCR conditions used. This experiment suggests there could be a mouse homologue of NAL in the library to be screened.

The Clontech mouse 11-day embryo 5'-STRETCH PLUS cDNA library was screened with the NAL probe which consisted of the full length clone. The probe was made by PCR of the NAL clone using primers M717 and M718, starting at base 13 and finishing at base 1143. The probe was then labelled with [32 P]dCTP as described in section 2.3.8. The primary screen produced 9 positive clones from a

total of a million screened. The inserts of three clones were then subcloned into the Bluescript vector so that they could be sequenced.

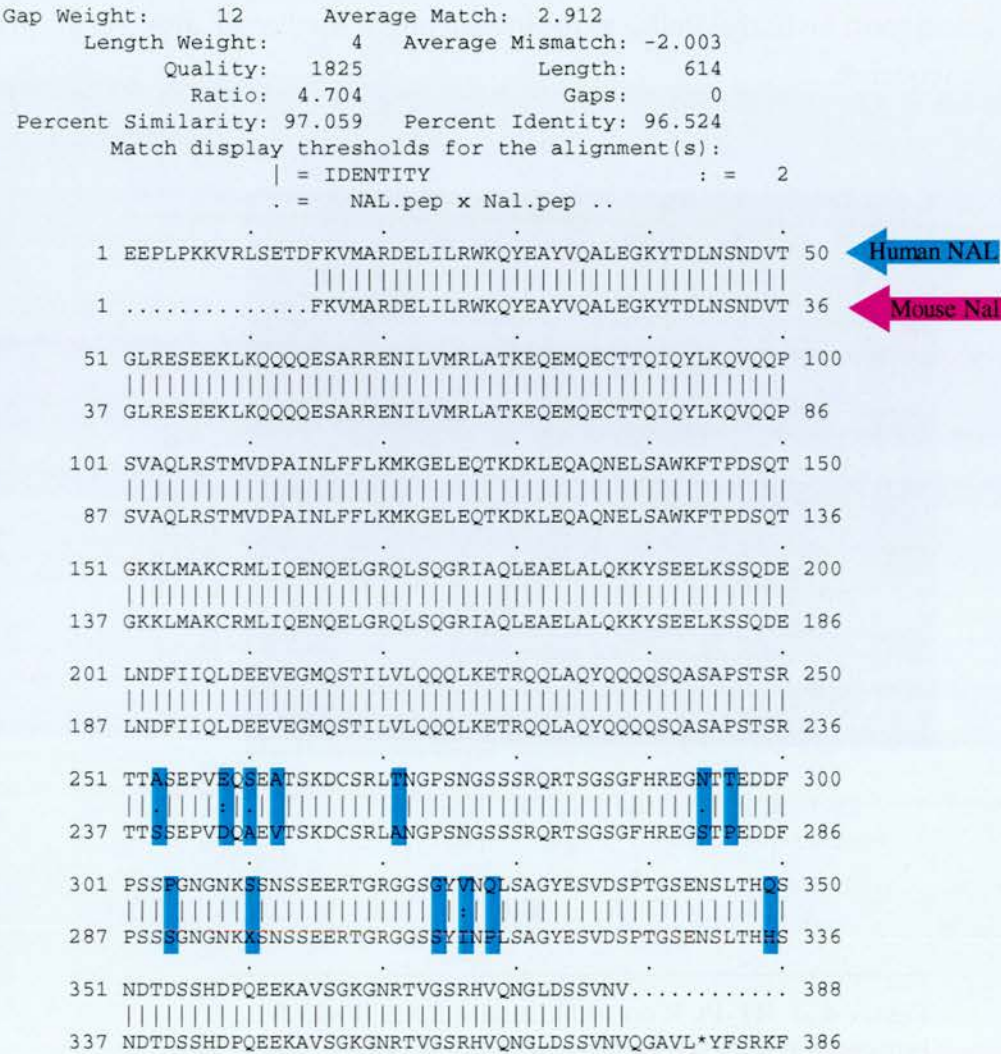


Figure 4.7 Comparison of the peptide sequence of the human NAL and mouse Nal clones. The two clones are 96% identical at the amino acid level, and the amino acids that differ are highlighted. The divergence occurs within a cluster and thus this region is probably not involved directly with the function of the protein. As can be seen, 6 amino acids after the end of the available human NAL sequence, the mouse NAL sequence has a stop codon, suggesting that the human NAL clone is not far off from a termination site.

4.4.2 The mouse and human clones of NAL are 94% identical

One of the mouse clones was sequenced fully and it was found that this mouse clone isolated matched NAL. At the DNA level the human and mouse NAL clones are 94% identical and at the protein level are 96% identical, as shown in figure 4.7. This shows that the mouse and human NAL are very highly conserved, and it suggests that the structure of the protein must be important for function as there is hardly any divergence. The fact that the mouse sequence is 94% identical to the human sequence also shows that NAL could not have originated from a chimaeric clone, as was thought at one stage (see section 4.2). Also a more recent search through the database showed that there was now a mouse EST clone which matched the one isolated, from position 744 to 1124. This shows that the library I had screened was not contaminated with human clones, as could be thought in light of the two clones being so similar, because the mouse clone in the database came from an independent source.

w26318.gb_est4 x NAL.seq ..

```

1 .....TTNNNNNTNCNCGGNCAGTGAAAACCTCTCTCACACACCAAT 41
      ::::: ::|||:|||||
1000 TGAAAGTGTAGACTCTCCACGGGCAGTGAAAACCTCTCTCACACACCAAT 1049
      42 CAAATGACACAGACTCCAGTCATGACCCCTCAAGAGGAGAAAGCAGTGAGT 91
      1050 CAAATGACACAGACTCCAGTCATGACCCCTCAAGAGGAGAAAGCAGTGAGT 1099
      92 GGGAAAGGTAATCGAACTGTGGGTTCGCGCCAGTTTCAAGATGGCTTGA 141
      |||||
1100 GGGAAAGGTAATCGAACTGTGGGTTCGCGCCAGTTTCAAGATGGCTTGA 1149
      142 CTCAAGTGTAATGTACAGGGTTCAGTTTGTGAAATATTTTTTCAGCAAAT 191
      |||||
1150 CTCAAGTGTAATGTACA..... 1167

```

Figure 4.8: Sequence comparison between the end of the NAL clone and a human clone from the database. This sequence comparison shows that the EST w26318, which originated from a human retina cDNA library, matches NAL, and the clone shows that there is a stop codon 5 amino acids after the end of the NAL clone.

4.5 The Human NAL clone is not quite full length

The human NAL clone which was isolated from the yeast two hybrid screen is not quite a full length clone. The clone does not have a start site, but by comparing its sequence to the human EST, D14661, it can be seen that it is only 3 amino acids away from the ATG site of the EST (see figure 4.2). The sequence around the ATG site has the conserved A three nucleotides upstream of the ATG, but the rest of the sequence is not a good Kozak consensus site (Kozak, 1984), so it is not definite that

this ATG is the start site. However, as discussed later in chapter 5, an antibody raised against NAL shows that the molecular weight of the endogenous protein is very similar to that of the NAL protein made *in vitro* from the NAL clone. The NAL clone also does not have a termination site, but there is a human clone in the database, originating from a retina library which matches NAL at position 1025, and goes through to a stop codon, as shown in figure 4.8. Also by comparing human NAL to mouse Nal, it can be seen that 6 amino acids after the end of the human clone sequence, the mouse clone has a termination site, followed by 3'UTR and a poly-adenylation site. Since the mouse and human clone are 96% identical at the amino acid level, it is likely that the human clone will be similar in length and so have a stop codon around the same place. The 3'UTR would, however, probably differ. If a full length NAL clone were to be needed, the rapid amplification of cDNA ends (RACE) technique could be used.

4.6 Mapping NAL

The EST D14661 had been isolated and placed in the database by a Japanese group that was trying to accumulate information on the structure of unidentified human genes by carrying out a big sequencing project of full-length cDNA clones. They therefore also mapped the EST to chromosome 6 (Nagase et al., 1995). However, the primers they used for mapping the EST were not documented and so it is not known if they were directed to the part of the sequence that did not match NAL, and because it is not yet totally clear whether this EST originated from a chimaeric clone, it was decided to map NAL independently, both in the human and mouse, by fluorescence *in situ* hybridisation (FISH) and also for the mouse by using the European collaborative interspecific backcross (EUCIB). By using these methods, the chromosomal location of NAL would also be more refined and it may provide some interesting findings if it maps to any disease locus.

4.6.1 Fluorescent *in situ* hybridisation

Mapping of the human NAL

Human NAL was mapped using FISH. For this purpose the full length NAL clone was digested out of the pGAD10 vector, and purified. FISH was carried out on human chromosome spreads and this was carried out by Muriel Lee. The FISH results (figure 4.9) showed that the human NAL gene mapped to chromosome 6, towards the telomere possibly between 6q25-27. This was encouraging because the EST also mapped to chromosome 6 (Nagase et al., 1995), increasing the likelihood that the EST is not from a chimaeric clone but is either an isoform of NAL or originates from unspliced RNA.

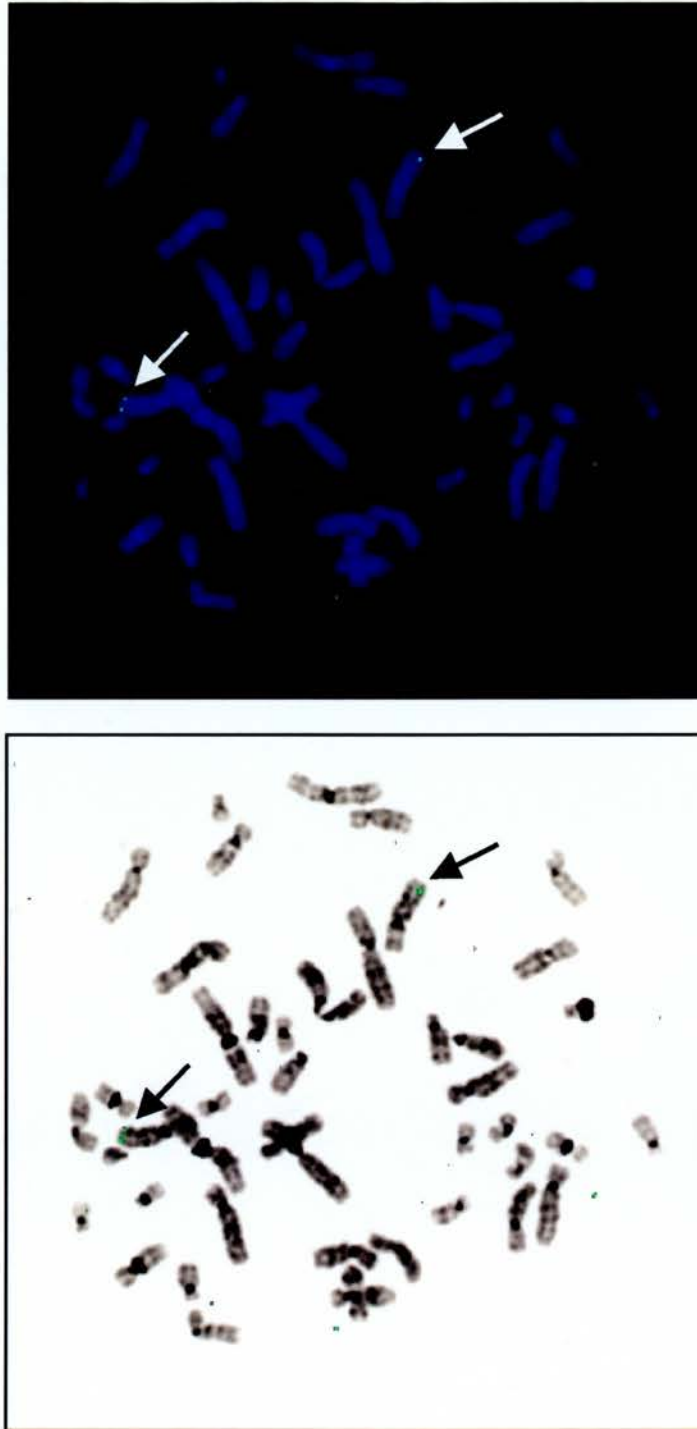


Figure 4.9: FISH analysis on human chromosome spreads using human NAL cDNA as the probe. The results show the signal appearing on chromosome 6, very close to the telomere. Some background signal can be seen, but this is because only a small cDNA probe was used. The FISH analysis was carried out by Muriel Lee.

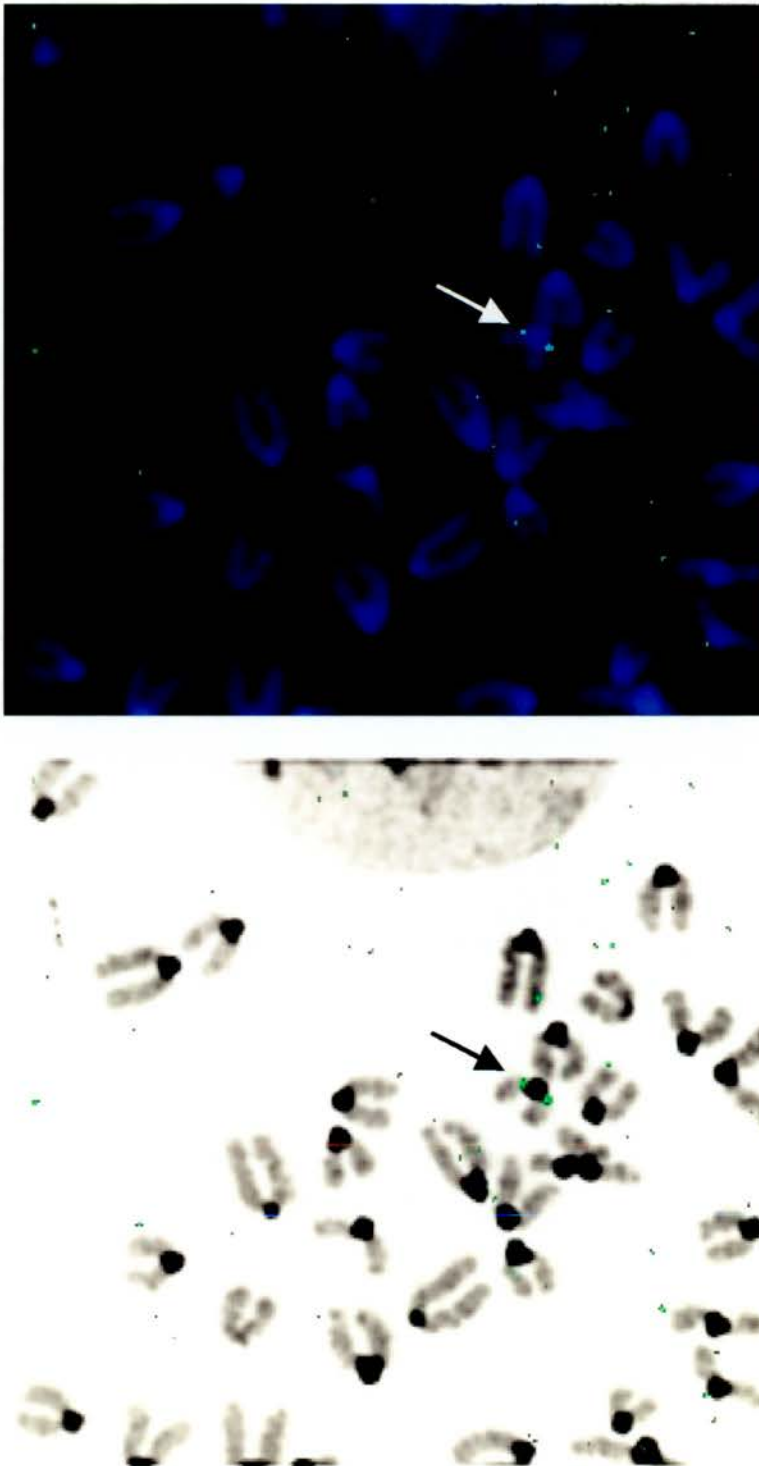


Figure 4.10: FISH analysis on mouse chromosome spreads using mouse Nal cDNA as the probe. The results show the signal appearing on chromosome 17, very close to the centromere. Some background signal can be seen, but this is because only a small cDNA probe was used. The FISH analysis was carried out by Muriel Lee.

Mapping of the mouse Nal clone

The mouse Nal gene was also to be mapped using FISH. The mouse insert was digested out of bluescript and purified. Again Muriel Lee carried out the FISH, this time on mouse chromosomal spreads. The results (figure 4.10) show that the mouse Nal gene maps to chromosome 17, near the centromere. This was also very encouraging as from linkage maps the mouse chromosome 17 and the human chromosome 6 show conserved synteny in the regions where the two genes map.

4.6.2 European collaborative interspecific mouse backcross

The EUCIB panel of mice provides a resource for genetically mapping a mouse gene at high resolution. It consists of 1,000 animal backcross progeny produced from a backcross between two mouse species, *Mus spretus* and C57BL/6 which was generated jointly between the UK HGMP centre and the French Institute Pasteur. F1 females are then backcrossed either to C57BL/6 or *Mus spretus*. The backcross progeny mouse are then scored for 3-4 markers per chromosome to produce an anchor map of 70 loci across the mouse genome (The European Backcross Collaborative Group, 1994). A new marker to be mapped first has to be linked to a chromosomal region and then analysed through a panel of mice identified as carrying recombinants within that chromosomal region. The new clone is then mapped to the closest marker that carries the same sequence of recombination events for each mouse tested.

Firstly, to detect whether a recombination event has occurred, a sequence variation within the clone of interest has to occur between the two species of mouse. Normally the sequence looked at is the 3'UTR, as this region of a gene is not so well conserved and therefore is more likely to have sequence variations. The 3'UTR of Nal was amplified in both strains of mice. However because in the laboratory there was no available DNA from *Mus spretus*, DNA from Ts/BL was amplified instead. The Ts/BL mice are a mixture of *spretus* and C57BL/6 mice, and thus the PCR should produce 2 bands: one corresponding to the C57BL/6 chromosome and one for the *spretus* chromosome. The PCR 5' primer started at position 1194 and the 3' primer started at position 1743 of the Nal clone, and this primer pair would amplify the 3'UTR. It was hoped that the 3'UTR would differ in length between the two sets of mice. Unfortunately no size difference could be resolved using PCR and so restriction digests were performed to find sequence polymorphisms between the two strains. It was found that one enzyme, DdeI, did show a difference between the two mouse strains. One band was produced for the C57BL/6 DNA digest and two bands

were produced for the Ts/BL DNA (one higher band representing the C57BL/6 DNA and the other representing *spretus* DNA), as shown in figure 4.11.

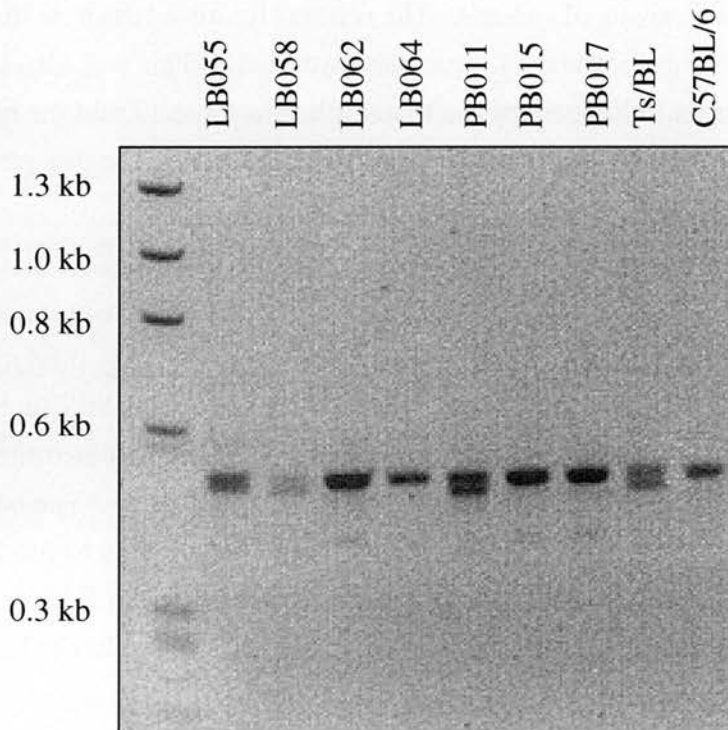


Figure 4.11: DdeI digests on test mice. The amplified PCR products for the 3'UTR of Nal from 7 test mice, were digested with DdeI and run on a 1.5% agarose gel. Test mice were chosen so that they had either both chromatids like the parental type or one chromatid of the parental type and the other chromatid from the other strain of mouse. In this way if the DdeI digest produced the expected number of bands, it could be used unambiguously as a diagnostic digest on the panel of mice used to map Nal. DNA from C57BL/6 mice and from Ts/BL (a mixture of *spretus* and C57BL/6) were run as internal controls.

To ensure that this PCR followed by a DdeI digest was a good way to distinguish between DNA from the two different mice strains, C57BL/6 and *Mus spretus*, the experiment was first tried on test mice. 7 mice were chosen in such a way that they had either both chromosomes as the parental type e.g. both C57BL/6, or had one chromosome from the parent and the other chromosome from the other mouse strain e.g. one from C57BL/6 and the other from *spretus*. The digests should then produce either one or two bands respectively. In this way the number of bands expected are

known, and thus it should be clear whether the diagnostic digest works without there being any ambiguity. Table 4.1 shows the backcross progeny used, together with the number of bands expected to be produced after a DdeI digest. The letters within the progeny name represent the strain of mouse to which the F1 generation were backcrossed to, as follows: LB= London Black 6; PB= Paris Black 6; LS= London *spretus*; PS= Paris *spretus*. Thus if an F1 is backcrossed to an LB and one band is produced, it suggests that for that locus, both chromosomes originated from C57BL/6 DNA. If the same backcross produces 2 bands, then one chromosome is like the parental C57BL/6 DNA whereas the other chromosome originates from *spretus* DNA.

Figure 4.11 shows the number of bands that were produced for each progeny after the DdeI digest, and these results are summarised in table 4.1. The results obtained show that the digests did produce the expected bands, and so the DdeI digest is a good way of distinguishing between the two mouse strains present.

Backcross Progeny	Expected number of bands	Number of bands produced by DdeI digest
LB055	2	2
LB058	2	2
LB062	1	1
LB064	1	1
PB011	2	2
PB015	1	1
PB017	1	1
Control C57BL/6	1	1
Control <i>Ts/BL</i>	2	2

Table 4.1: DdeI digest of the 3'UTR of Nal of test mice. The table shows the mouse progeny used to test whether the DdeI diagnostic digest produces the expected results. This experiment showed that the DdeI digest can distinguish between the DNA originating from C57BL/6 mice and the DNA originating from *spretus* mice.

The diagnostic digests could now be carried out on a larger panel of mice so as to be able to map Nal. The first stage of the mapping procedure - linking Nal to a chromosomal region - had effectively already been done by using FISH. Therefore

rather than having to take a random panel of mice to achieve this chromosomal linkage, mice were chosen according to whether they would be informative. This meant choosing mice that had as many markers mapped to the candidate Nal region. These mice were also chosen according to whether their markers had been ordered on the chromosome via one recombination event only. This is because usually only one recombination event occurs and so if markers are mapped onto a chromosomal region by the formation of 2-3 recombination events, it suggests that perhaps a typing error had occurred. The informative mice were those that had markers on chromosome 17 from the centromere up to marker D17Mit9. This region was the selected region because from the mouse FISH data, Nal is known to map to chromosome 17 near the centromere. The human FISH data shows that NAL maps to chromosome 6q near the telomere. When a linkage map of chromosome 17 is looked at, it can be seen that the region between the centromere and marker D17Mit9 shows conserved synteny with the telomeric end of human chromosome 6, and so it is likely that Nal maps in that region.

The MBx database, which supports the backcross, was used to find the panel of mice that had markers within the chosen region of chromosome 17. From this panel, 45 mice were chosen for the mapping so as to make the experiment statistically significant. The 3'UTR of Nal was amplified from these mice. Only 38 tubes of DNA from the backcross progeny amplified the product and this was because a lot of the tubes were old and so contained no DNA. The backcross progeny used and the haplotype for each marker, as determined by EUCIB, are shown in table 4.2. The PCR products that had been amplified were then digested with DdeI and the number of bands produced were recorded, as shown in table 4.3. The number of bands produced for each mouse helps to pinpoint the Nal locus, as the markers in that region which are linked to Nal will have the same number of bands, as shown in table 4.2. By using a large number of mice, the markers which are closely linked to Nal will become apparent. Table 4.2 highlights in grey those markers which are linked to Nal for each backcross progeny.

These results map Nal at least as far up as marker D17Mit25, because 35 out of 38 recombinants are concordant with this anchor marker. In other words, each backcross progeny used to map Nal, has produced a series of 1 or 2 bands, and all 35 of them correspond to the number of bands which have been produced for the marker D17Mit25. This suggests that Nal is very closely linked to this anchor marker because the probability of getting the same output of ones and twos for a marker that Nal does not map close to is minimal. There is a possibility that Nal maps even closer to the centromere than marker D17Mit25, because one of the mice which was not concordant with the anchor marker was mouse PB050. The linkage seen for this mouse would place Nal up as far as marker D17Mit97, and the limited data available for this marker still agrees with the output produced for Nal. The other two markers which were not concordant with the anchor marker D17Mit25, were mice LS583 and LS-F780. However, because both of these mice have no data for markers just above D17Mit25 (table 4.2, highlighted in dark grey), it could be that another recombination event could have occurred in these mice and this would place Nal just above D17Mit25.

According to the linkage map produced for chromosome 17 on the Jackson Laboratory mouse genome informatics web page, marker D17Mit25 is 4cM from the centromere, and therefore Nal maps centromeric to the Brachyury (T) and within the T-locus of the mouse.

Backcross Progeny	No. of Bands After DdeI Digest	Site Mapped To (refer to table 4.2)
LB202	1	Above Mit43
LS285	2	Above Mit102
LS215	2	Above Mit174
PB044	2	Above Mit108
LS554	2	Above Mit105
PS044	1	Above Mit197
LS318	2	Above Mit105
PB050	1	Above Mit97
LS226	2	Above Mit51
LS276	2	Above Mit55
LS-F772	1	Above Mit225
PB202	1	Above Mit62
LS-F786	1	Above Mit51
LS-F787	1	Above Mit11
LS556	1	Above Mit216
LB162	2	Above Mit80
LB-F411	1	Above Mit25
PS028	1	Above Mit68
PB191	1	Above Mit108
PB186	1	Above Mit108
PS031	1	Above Mit68
LS181	2	Above Mit106
LS586	1	Above Mit225
PB041	2	Above Mit51
LS700	1	Above Mit225
LB237	2	Above Mit51
LS695	2	Above Mit85
LS690	1	Above Mit25
LB252	1	Above Mit102
LB354	2	Above Mit46
LB356	1	Above Mit22
LS174	1	Above Mit68
LB201	2	Above Mit125
LS-F771	2	Above Mit167
LS583	2	Below Mit21
LS-F777	1	Above Mit21
LS-F780	1	Below Mit60
PS049	2	Above Mit51
Control CB57BL/6	1	-
Control Ts/BL	2	-

Table 4.3: Mapping results for Nal using EUCIB mice.

4.7 Conclusions

NAL which binds to WT1 in the yeast two hybrid system, has been shown to be a novel protein which has not been characterised before. The database searches showed that NAL matched an EST, D14661, but only across the first 445 bases, after which there was a breakpoint where all homology was lost. This was intriguing and suggested that maybe one of the two originated from a chimaeric clone. It was important to determine whether NAL was expressed *in vivo* by carrying out RT-PCR experiments. The results showed that NAL was not chimaeric and pointed towards the EST in the database being composed of two non-related cDNAs. However further analysis suggests that perhaps the EST is an alternative transcript, although Northern blots would have to be carried out to confirm this. The Northern blots would also help determine whether the transcripts are tissue specific.

A mouse homologue of NAL was isolated from a library screen and shown to be 94% identical at the DNA level, and 96% identical at the protein level, to human NAL. This sequence conservation suggests that the structure of the protein must be very important for function as there is very little variation. No other homologues have been identified in the database, not even the fungi or invertebrate database, suggesting that NAL is only found in vertebrates. The only known motif that is present in the protein is a stretch of polyglutamines. This motif is often associated with transcription factors, and so further work could be done to assess whether NAL plays a role in transcription.

Mapping studies using FISH on mouse chromosomes as well as segregation analysis in the EUCIB resource, have mapped mouse Nal to chromosome 17, around 4cM from the centromere. This region of chromosome 17 is home to the T-locus (Bennett, 1975). The T-locus in the mouse is a region of chromosome identified by sets of dominant and recessive mutations, some of which have profound effects on embryonal development, sperm production and function and genetic recombination in a large region of their chromosome. The origin of the T-locus region began with the detection of the dominant mutation, *Brachyury* (T), which produces short tails in heterozygous mice but is lethal in homozygotes. Recessive mutations produce normal tailed heterozygotes, and they interact with dominant T-alleles to form a tailless phenotype. Recessive t-alleles of independent origin can complement each other as they represent different mutations. Different t-alleles produce different embryonic abnormalities when they are homozygous, and examples include *fused* which influences tail length; *quaking* which is a neurological disorder; and *tf* which causes hair loss. This suggests that this region of chromosome 17 has several genes

which are essential for normal development. It is not really known how much of chromosome 17 the T-locus occupies, but it probably spreads for about 16cM. No t-alleles have been isolated so close to the centromere where Nal maps and so it is not known therefore whether Nal is part of this T-locus and whether it is one of the genes essential for development.

The human NAL gene has been mapped to chromosome 6q26-27 by FISH on human chromosomes. This region of the human chromosome 6 shows conserved synteny with the centromeric region of the mouse chromosome 17. Interestingly 6q26-27, has been associated with several malignancies and therefore it is thought that a tumour suppressor gene may lie in the region. These malignancies include renal cell carcinoma (Morita et al., 1991); ovarian cancer (Cooke et al., 1996); and acute lymphoblastic leukaemia (Menasce et al., 1994). It is interesting to note that mutations in WT1 have been found associated with some of these diseases too, as explained in section 1.8. It could be envisaged that if WT1 and NAL do interact, then mutations in both genes could lead to multistep carcinogenesis. However, it has not been shown that NAL is the gene in 6q which predisposes to any of these malignancies. It would be of interest to get some tumour tissue to see if the structure of the gene and the expression of NAL is altered compared to normal tissue.

Chapter 5

Binding Studies

Chapter 5: Binding Studies

The yeast two hybrid system identified NAL, a novel protein, as potentially binding to WT1, as discussed in chapter 3. Although within the Y2H system the interaction between the two proteins appeared real and specific, other binding assays had to be carried out to confirm that the interaction is physiologically relevant. This is because, despite the yeast two hybrid system being a powerful tool used to identify protein-protein interactions, both proteins are overexpressed in the yeast and therefore there is a possibility that they will be forced to interact, where under normal circumstances they might not.

5.1 *In Vitro* Binding Assays

In vitro binding assays are a good way to determine whether two proteins interact. Although these assays use excess protein and therefore, like the yeast two hybrid system, do not provide evidence that the two proteins will interact at their endogenous levels, they do provide an alternative system to test the interaction. Thus if the interaction between NAL and WT1 still occurs in these *in vitro* assays, then it will indicate that the two proteins can bind to each other in systems other than the Y2H.

5.1.1 Problems Using Bacterially Expressed pET6HWT1

To test whether NAL can bind to WT1 *in vitro*, bacterially expressed WT1 and *in vitro* translated NAL were to be used. The idea was to clone WT1 into the pET6H expression vector, which would fuse 6 histidine residues to the N-terminus of WT1. The 6 histidine residues would allow the bacterially expressed 6H-WT1 protein to be purified over a nickel-chelate affinity resin. An *in vitro* transcription/translation kit (Promega, section 2.6.1) would be used to express [³⁵S]NAL. This labelled protein would then be mixed with 6H-WT1 before running over the nickel column. The 6H-WT1 would be purified from the mixture, and if NAL binds to WT1 it should be co-purified, and not be removed during the washing steps. The presence of any bound [³⁵S]NAL would then be detected using fluorography (see section 2.6.8).

To express 6H-WT1, the WT1 was cloned into the pET6H vector (this had previously been carried out by Dr. Davies) which contains the 6 histidine residues just after the translation initiation site. The vector can then be transformed into bacteria and the expression of 6H-WT1 induced by the addition of IPTG, as discussed in section 2.6.2.

Time Course Studies

To express the 6H-WT1 protein in bacteria, the expression vector had to be induced with IPTG. The time of induction had to be optimised for the best expression levels of the protein. For this time course experiment, the vector was transformed into bacteria which were left to grow at 30°C to an OD₆₀₀ of 0.7. Then before 1mM IPTG was added, a sample of bacteria was taken, and this was to be time point 0. The bacteria were then induced and samples taken at time points corresponding to 30 minutes; 1 hour; 2 hours; and 3 hours (see materials and methods). The bacterial samples at the different time points were then resolved on a 10% SDS -PAGE to see if there was any expression of 6H-WT1 - see figure 5.1.

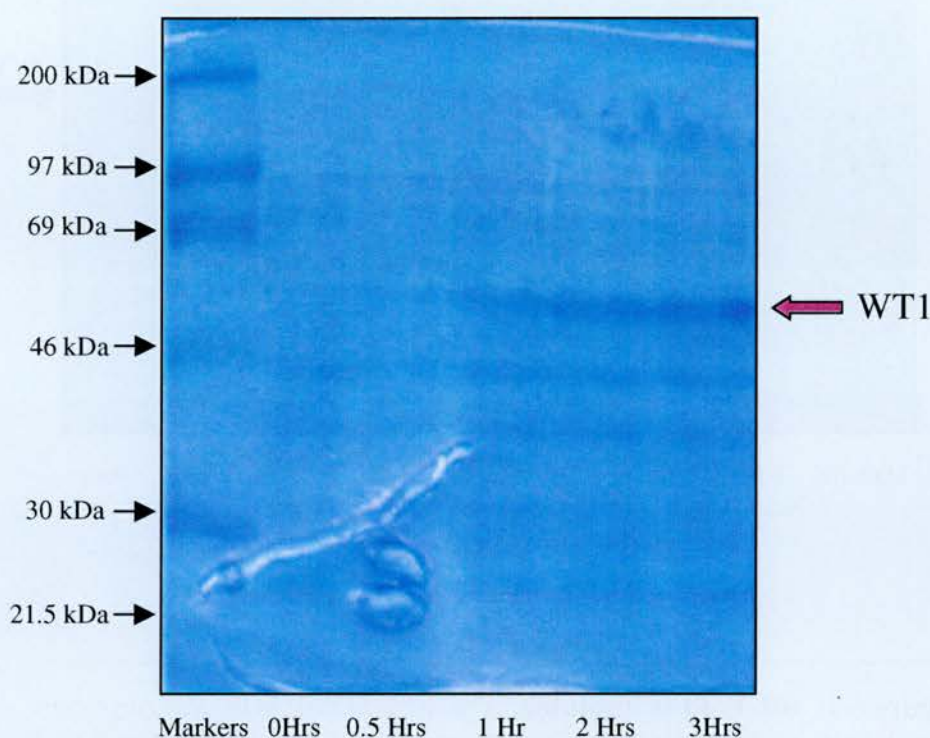


Figure 5.1: Induction time course for pET6H-WT1.

The 10% SDS PAGE shows that high levels of 6H-WT1 are induced after 2-3 hours.

The gel showed that expression of WT1 had occurred after induction with IPTG, as compared to the pre-induced sample. Higher levels of expression were seen 2 and 3 hours after the start of induction. The 1 hour sample did express 6H-WT1 but to a lower level. Thus for the rest of the experiments, the induction time ranged between 2-3 hours.

6H-WT1 is insoluble in bacteria

The 6H-WT1 protein expressed in bacteria had to be soluble so that it could be purified and used in the *in vitro* studies. Thus after the induction, the bacterial cultures were lysed as discussed in section 2.6.2 and soluble proteins separated from insoluble ones by centrifugation. Samples at each stage of the procedure were resolved on 10% SDS PAGE. All four isoforms of WT1 had been expressed in bacteria so as to be able to perform comparative binding assays.

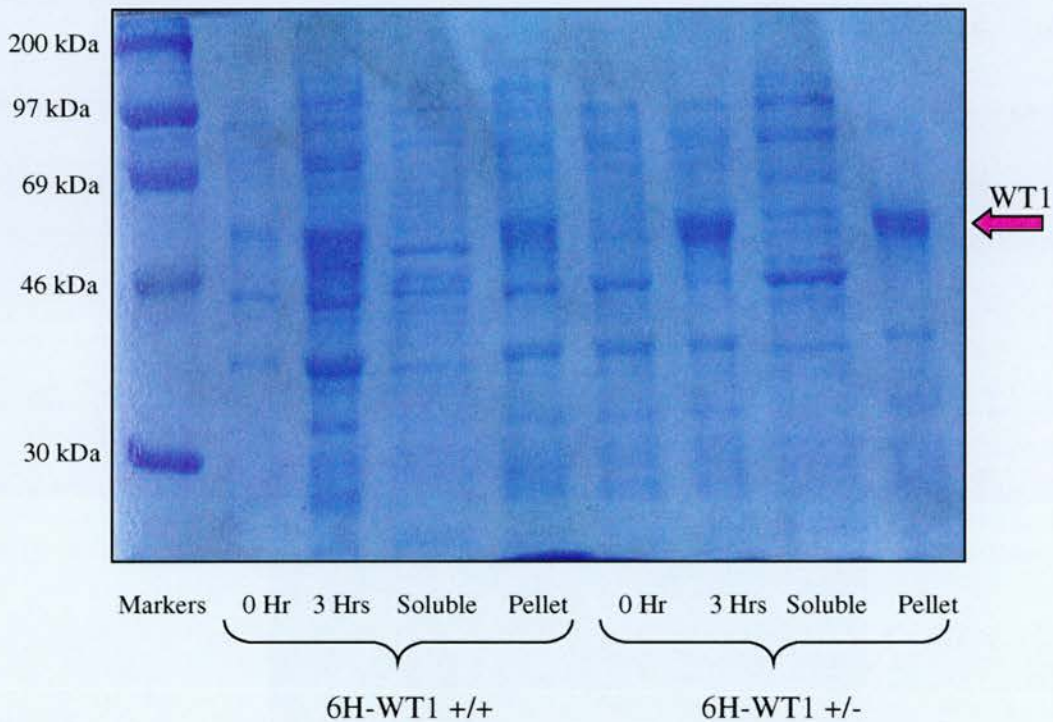


Figure 5.2: 6H-WT1 is insoluble. The 10% SDS-PAGE was run from samples taken for 6H-WT1 +/+ and +/- expressed in bacteria. The 0 hr and 3 hrs samples represent the time points after induction with IPTG, and are whole bacterial lysates. After 3 hours of induction the bacteria were lysed and both the soluble and insoluble fractions were run on the gel. As can be seen for both 6H-WT1+/+ and +/- the protein was expressed after 3 hours, but was found in the pellet fraction as it was insoluble.

The gel in figure 5.2, which only presents the data for WT1+/+ and +/- isoforms, shows that unfortunately the 6H-WT1 is predominantly insoluble as most of the protein is found in the pellet fraction and not into the soluble fraction. Rauscher *et al.* had also had problems when they tried to use 6H-zinc fingers of WT1, as they

were not soluble and so had to be re-suspended in 6M guanidine-HCl (Rauscher et al., 1990). Although potentially the 6H-WT1 protein could be solubilized in this way, it was decided against as the binding properties may be altered following the denaturation/renaturation procedure. Instead, in an attempt to solubilise the protein, the induction temperature was lowered to 25°C, as the different conditions can sometimes help with solubility. Unfortunately at this temperature the protein was not expressed. It was therefore decided to use the glutathione S-transferase (GST) expression system in the hope that in this way WT1 may be more soluble.

5.1.2 Problems Using Bacterially Expressed GST-WT1 Zinc Fingers

The pGEX glutathione S-transferase (GST) fusion protein system can be used to express and purify high levels of fusion proteins from bacteria. The 26kDa GST enzyme binds reversibly and with high affinity to glutathione-containing affinity matrices. Thus the soluble fraction of bacterial lysates, containing the GST fusion protein, can be incubated with glutathione agarose beads and the fusion protein purified away from other bacterial proteins. The fusion protein can then be eluted from the beads by using glutathione elution buffer.

As seen from the previous experiment, 6H-WT1 is fairly insoluble. Therefore rather than use full length WT1, only the zinc fingers were to be used as these had been soluble in the past (Bickmore et al., 1992). It was also known that NAL bound to the zinc finger of WT1 (section 3.5.2) and so the *in vitro* binding studies could be performed with just the zinc fingers of WT1.

Section 2.6.3 describes the expression of the GST-WT1 zinc fingers in bacteria. An induction time course was performed, as was done for 6H-WT1, but this time the induction was carried out at 25°C, and it was found that again a good induction period ranged between 2-3 hours. After a 3 hour induction, the bacterial cultures were lysed as described in section 2.6.4. Both the soluble and pellet fractions were run on a 10% SDS PAGE to analyse whether the protein was soluble.

The gel in figure 5.3 shows that unfortunately GST-WT1 zinc fingers were insoluble, as the protein was found in the pellet fraction. There may have been a small amount of the protein in the soluble fraction, which had less sample loaded onto the gel compared to the other samples, but in comparison to the amount of protein required for the binding assays, it was not nearly enough. The induction time and temperature were altered to try and get more soluble GST-zinc fingers but the fusion protein remained insoluble. Different host bacterial strains were also used, but the protein was still insoluble. At this stage it was decided that rather than persevere

with trying to solubilise WT1, it was best to switch the experiment round so that NAL became fused to the GST, in the hope that this fusion protein would be more soluble than WT1.

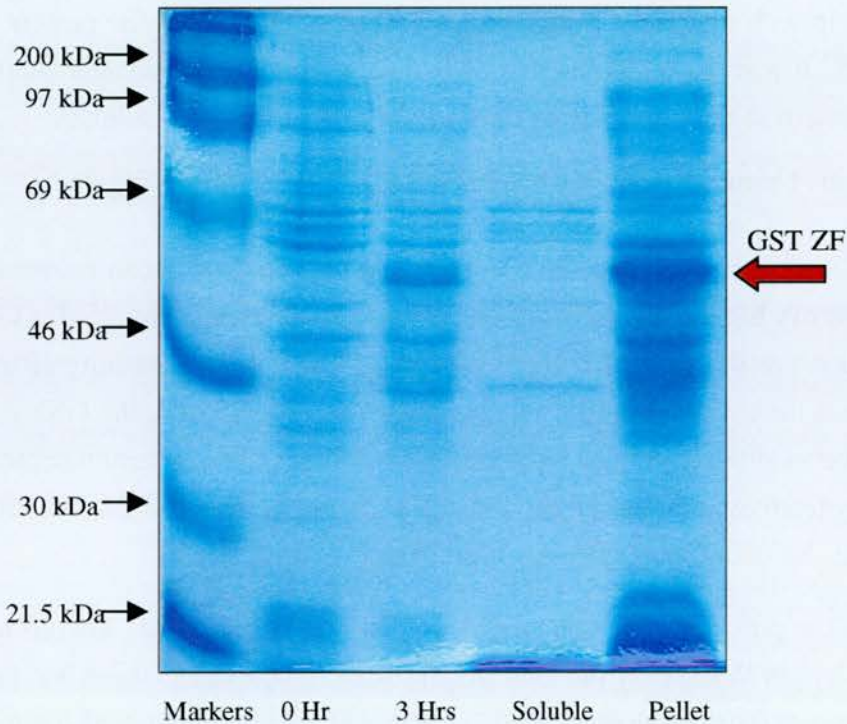


Figure 5.3: Bacterial expression of GST- WT1 Zinc Fingers. The 10% SDS-PAGE shows that the GST-WT1 zinc fingers (-KTS) were expressed after 3 hours of induction. However, unfortunately the fusion protein is predominantly insoluble as it is found in the pellet fraction.

5.1.3 Using Bacterially Expressed GST-NAL

NAL first had to be cloned into the pGEX-5X-1 vector. To do this NAL was digested out from the pGad10-NAL vector, which had been used in the Y2H system, as an EcoRI fragment and cloned straight into the pGEX vector.

The GEX-NAL vector was then transformed into bacteria, expression induced and bacterial cultures lysed as for the GST-zinc fingers. An SDS-PAGE was run to see if GST-NAL was soluble. This time it could be seen that there was a large amount of GST-NAL protein which was soluble despite some of the protein still being present in the pellet fraction, as is shown in figure 5.4. Nevertheless there was sufficient

protein in the soluble fraction which could then be purified as detailed in section 2.6.4.

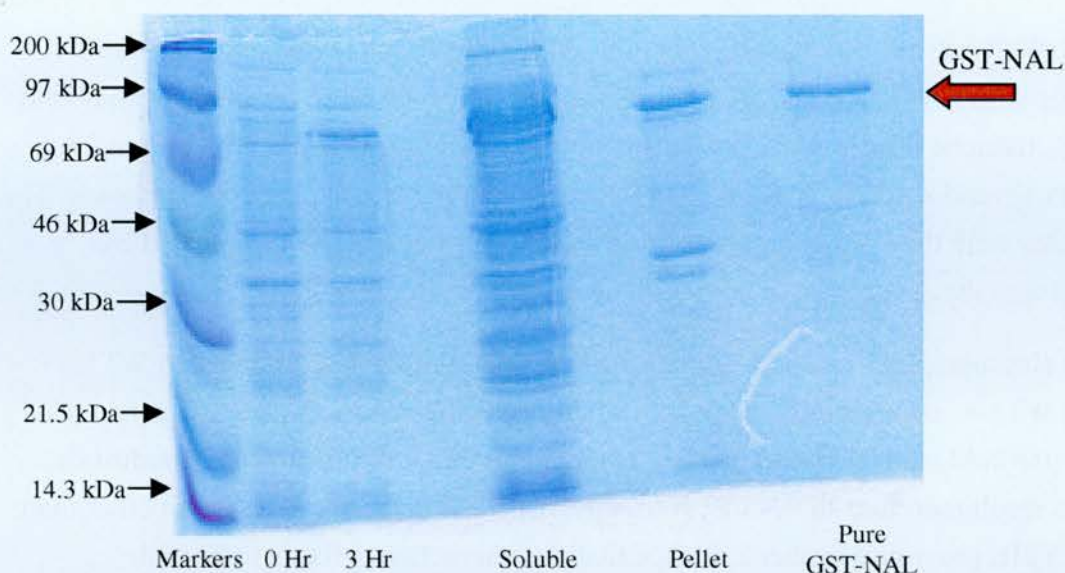


Figure 5.4: GST-NAL is partially soluble. The 10% SDS-PAGE shows that the GST-NAL fusion protein was induced after 3 hours. The soluble fraction does contain GST-NAL, even though it is not all soluble as the pellet fraction still contains some of the fusion protein. However the soluble fraction contains enough of the protein and, as the last lane of the gel shows, the GST-NAL fusion protein could be purified.

Binding Assays Under Low Salt Conditions

The GST-NAL could now be used in the *in vitro* binding assays together with *in vitro* translated WT1. In these binding assays, GST-NAL and *in vitro* translated [³⁵S]WT1 are allowed to mix under low stringency conditions, by using a low salt concentration of 100mM. The mixture is then run through glutathione agarose beads, from which GST-NAL is purified. If WT1 binds to NAL then it should be co-purified, rather than being removed during the washing steps. To detect the presence of WT1, the purified complex is run on an SDS PAGE to resolve the two proteins, and WT1 can then be detected by fluorography as it is radioactively labelled. For these binding assays, a control has to be run to ensure that WT1 is binding to NAL and not to the GST moiety or the beads. Therefore a parallel experiment is carried in exactly the same manner, except that this time instead of using GST-NAL, the GST protein on its own is mixed with [³⁵S]WT1. In this

negative control, no WT1 should be detected on the fluorograph unless WT1 is being purified non-specifically.

[³⁵S]WT1 was prepared as described in section 2.6.1 and it was allowed to mix either with GST-NAL or GST on its own, both of which had previously been bound to glutathione beads, under low stringency conditions as described in section 2.6.6. After several washing steps, the GST and GST-NAL were eluted from the beads. The eluates were then run on an SDS PAGE and the presence of WT1 detected by fluorography.

The fluorograph produced (figure 5.5) shows that both *in vitro* translated WT1+/+ and WT1-/- could bind to GST-NAL but not to GST alone. This demonstrates that *in vitro* NAL and WT1 can bind together, albeit under low stringency conditions. This result therefore shows that NAL and WT1 can associate in a system other than the Y2H, providing further evidence that the interaction is likely to be real.

Binding Assays Under Higher Salt Concentration

To show that the interaction between NAL and WT1 can occur under higher stringency conditions, the salt concentration was increased from 100mM to 250mM. These higher stringency experiments would also include all four WT1 isoforms and deletion constructs described in section 3.4.4 to test the specificity of the interaction between NAL and the C-terminus of WT1 and to test whether NAL could bind to all four isoforms of WT1. As the fluorographs show, all four isoforms of WT1 can bind to NAL under these higher salt concentrations (figure 5.6, panels A and B), suggesting that the interaction between WT1 and NAL is reasonably strong. However these assays proved insufficiently sensitive to detect any subtle differences in NAL binding affinity between the WT1 isoforms, even when the salt concentration was increased to 500mM (data not shown).

This *in vitro* assay was able to show that NAL can bind to the deletion constructs Nco3'+KTS and -KTS but cannot bind to the deletion construct Nco5' as shown in figure 5.6, panels C and D. The interaction between WT1 and NAL is shown to be specific as NAL hardly binds to EGR1, another zinc finger protein (figure 5.6, panel D).

The results obtained in these *in vitro* assays were very encouraging as they provide further evidence that NAL is a WT1-binding protein. The Y2H data and the *in vitro* data both agree that NAL binds to the C-terminus of WT1, although no isoform specific differences were apparent in the *in vitro* experiments.

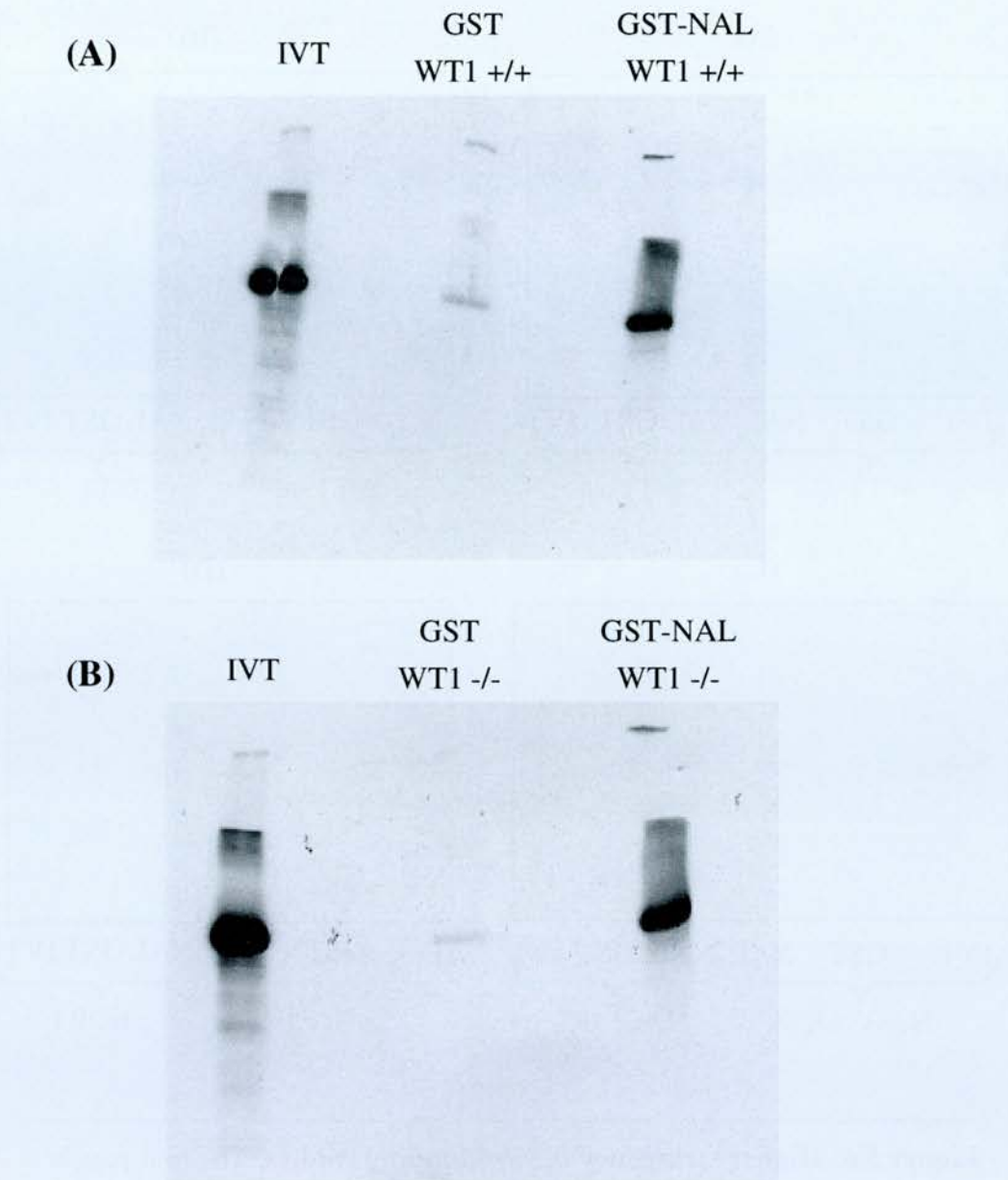


Figure 5.5: *In vitro* binding studies between WT1 and NAL.

The fluorograph in panel (A) shows in lane 1 an *in vitro* translation (IVT) check to show that [^{35}S]WT1 was made. The band in the last lane, labelled GST-NAL/WT1+/, represents WT1+/+ which was co-purified with the GST-NAL, indicating that WT1+/+ can bind to NAL *in vitro* under low salt conditions. The middle lane, labelled GST/WT1+/, indicates that there is very little residual binding of WT1 to GST alone. Panel (B) is exactly the same as panel (A) but this time WT1-/- was used.

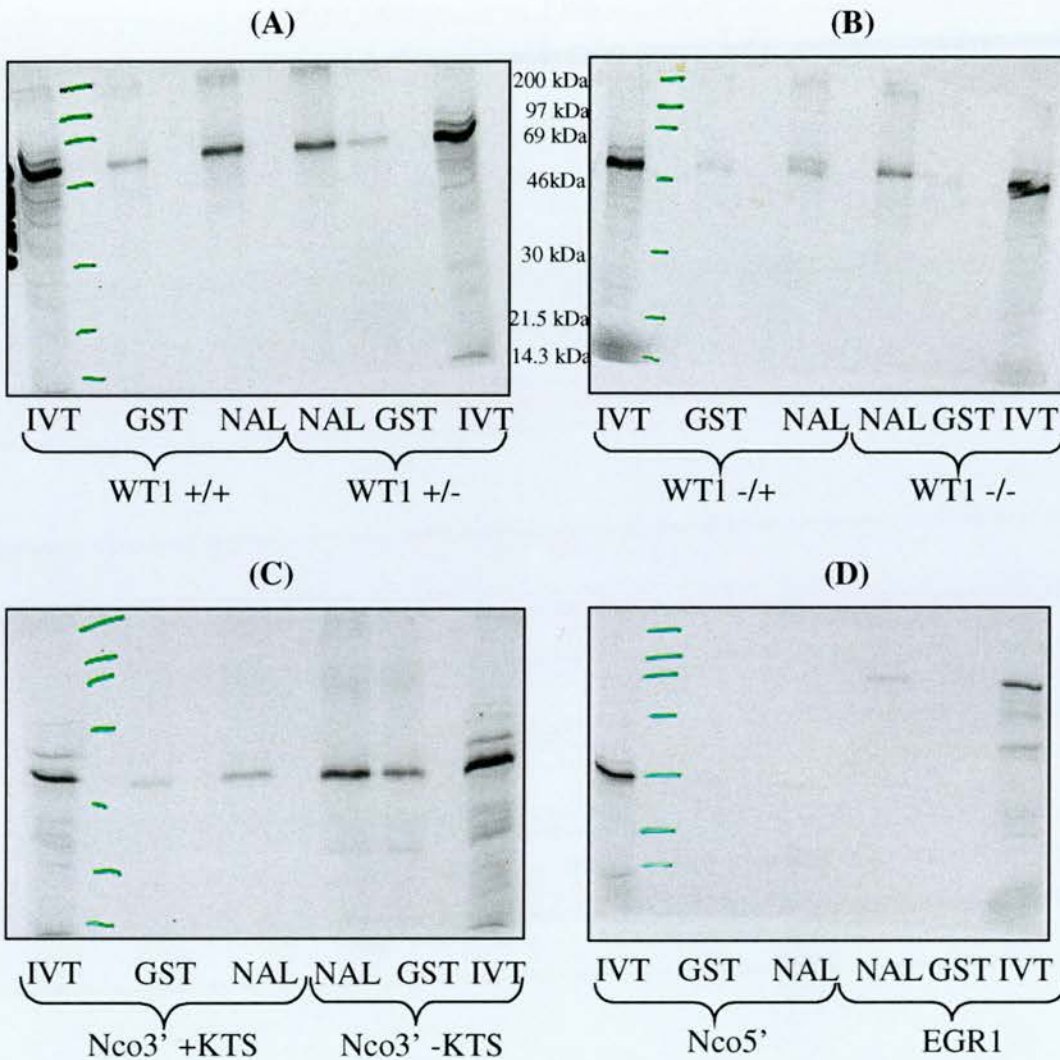


Figure 5.6: Higher stringency *in vitro* binding studies. The four panels represent the results produced for the binding studies carried out under higher salt conditions (250mM salt) between GST-NAL, as well as the GST controls, and *in vitro* translated WT1. Each panel represents two experiments, each comprising of an *in vitro* translation check (IVT), the binding seen with GST on its own, and the binding seen with NAL. The WT1 isoform or deletion construct used is shown under the bracket. Panel (A) shows that WT1 +/+ and WT1 +/- can bind to NAL but not to GST only, which only shows slight background bands. Panel (B) shows that WT1 -/+ and WT1 -/- can also bind to NAL. Panel (C) shows that Nco3' +KTS and Nco3' -KTS can bind to NAL, although the background levels for GST interacting with Nco3' -KTS are also rather high. Panel (D) shows that the Nco5' deletion construct cannot bind to NAL, as was also shown in the yeast two hybrid system. A very slight band can be seen for the interaction with EGR1, suggesting that perhaps NAL can bind to EGR1, but the interaction does not seem to be as strong as that seen for WT1.

5.2 *In Vivo* Assays Using HA Epitope-Tagged NAL

The *in vitro* data, like the yeast two hybrid data, shows that NAL can bind to WT1. However, these assays overexpress both proteins and are not carried out under endogenous conditions and so may not represent the *in vivo* situation. Therefore to ensure that the interaction was physiologically relevant, *in vivo* binding assays were carried out.

Ideally these *in vivo* assays would be carried out on cells that endogenously express both proteins. However, these *in vivo* assays require both proteins to be detected by antibodies, and as NAL is a novel protein there are no antibodies available against it. To overcome this problem, NAL could be epitope tagged, transfected into WT1 expressing M15 cells, and detected by using an antibody against the epitope. Co-immunoprecipitation experiments can then be carried out as detailed in section 2.8.8. Essentially nuclear extracts from these transfected cells are made, and WT1 is immunoprecipitated from these extracts by incubating them with a WT1 antibody, which itself has been crosslinked to protein A sepharose beads. WT1 is then purified from the lysate, together with any proteins that bind to it. The WT1 complex can then be resolved on a SDS-PAGE, and an immunoblot carried out as described in section 2.6.9. Thus to determine whether NAL binds to WT1, the immunoblot is probed with an antibody raised against the epitope with which NAL has been tagged. In this way it should be possible to establish whether NAL and WT1 form a complex *in vivo*.

5.2.1 Epitope Tagging NAL

The epitope tag to be used was a haemagglutinin (HA) motif which the monoclonal antibody 12CA5 can detect. A pCI-neo vector (Promega), which had three N-terminal HA motifs engineered into it (gift from Matteo Ruggiu), was used to epitope tag NAL. NAL was cloned into this vector as a XbaI/SalI fragment, these sites having been added to the NAL clone by PCR using primers N842 and N845 (see appendix).

5.2.2 Testing Expression Of The HA-NAL Clone

Before the binding assay could be done, the HA-NAL clone had to be transiently transfected into the M15 cell line, and preliminary studies carried out to ensure the tagged construct could be easily transfected, expressed and detected in these cells. HA-NAL was thus transiently transfected into the M15 cell line, as detailed in section 2.8.6. As a control, the M15 cells were also transfected with a CMV vector

which did not have the HA tag or NAL cloned into it. The cells were lysed 24 or 48 hours after transfection and equal amounts of each lysate run on an SDS PAGE. An immunoblot was carried with the HA antibody, 12CA5, to detect the HA-NAL clone. Figure 5.7 shows that the HA-NAL protein was detected by the 12CA5 antibody, and therefore was properly transfected, expressed and detected. The better expression levels of the HA-NAL protein were seen 48 hours after transfection. The HA antibody did not detect anything in the negative controls which had the CMV vector transfected in the cells.

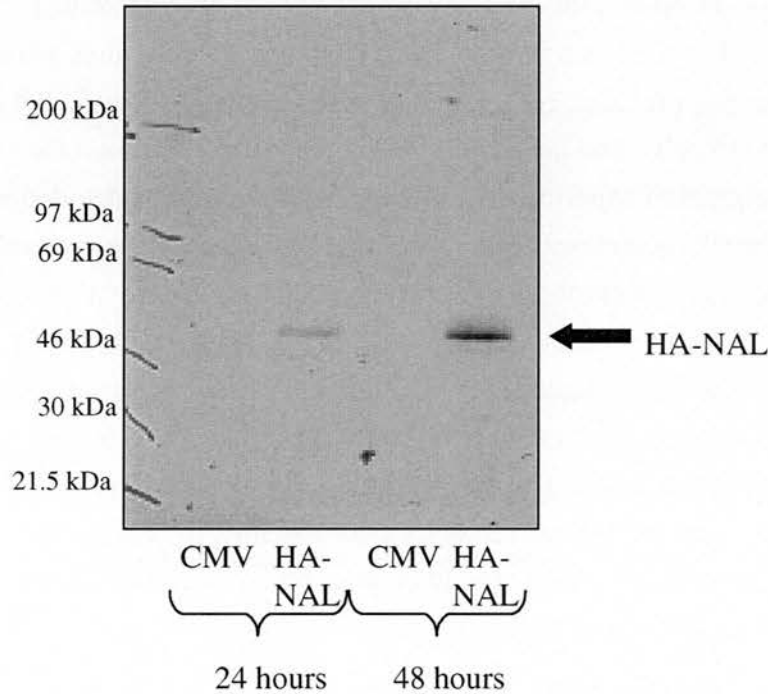


Figure 5.7: Expression studies for transiently transfected HA-NAL. The immunoblot shows the detection of HA-NAL in transiently transfected M15 cells. The expression levels is greater after 48 hours compared to the expression level seen 24 hours after transfection. The negative controls, where a CMV vector containing no HA-NAL was transfected in M15 cells, show no detection of the HA-NAL protein.

5.2.3 Binding Assays With HA-NAL

Now that the conditions for transfecting and detecting HA-NAL were set up, the co-immunoprecipitation assays could be carried out. The cells were transfected with HA-NAL as before and left for 48 hours to get high levels of HA-NAL expression. Using a β -galactosidase vector as a control for the efficiency of the transfection (see section 2.8.7) it was shown that the transient transfection had worked. After this time, the cells were harvested, and nuclear extracts made, as detailed in section

2.8.8. The nuclear extracts were then split into three, and mixed with either a WT1 antibody, H2, to carry out the binding assay; the HA antibody, 12CA5, as a positive control; or mouse IgG, as a negative control. The immunoprecipitations were performed, as is detailed in section 2.8.8, and resolved on an 8% SDS PAGE. An immunoblot was then carried out with the 12CA5 antibody to detect the presence of HA-NAL in any of the immunoprecipitates.

Figure 5.8 shows the results obtained for the binding assay. The whole cell extract (WCE) lane had been run as a control, to demonstrate that the cells that had been transfected with HA-NAL were indeed expressing it. This lane was also run as a positive control to ensure the detection step of the experiment was working. The positive control lane that contained the immunoprecipitate from the 12CA5 also showed a band as expected. This positive control used the 12CA5 antibody to immunoprecipitate HA-NAL from the cells, and so shows that the experimental conditions were right for the immunoprecipitation. The lane which contained the WT1 immunoprecipitate, and was the lane which was to show whether NAL and WT1 interact, also had a band. Unfortunately this band probably does not represent HA-NAL, but instead the heavy chain of some WT1 antibody which had not been completely crosslinked to the sepharose beads prior to the start of the experiment. The heavy chain of the antibody is of a very similar size to the HA-NAL protein and unless it is covalently crosslinked to the sepharose beads it would run at a similar position to HA-NAL on an SDS-PAGE. If there is any antibody which remains uncrosslinked, then it would be recognised by the secondary antibody during the immunoblot procedure thus producing a band which does not represent HA-NAL. Normally any antibody which is not crosslinked is removed by the washing steps carried out with glycine after the crosslinking procedure. However, in this particular experiment it seems that the washes with glycine were not thorough enough as there was still some residual antibody which had not been crosslinked.

Although the heavy chain of the antibody and HA-NAL run very close together they could be distinguished on this 8% SDS PAGE. If the HA-NAL band produced in the WCE lane is compared to the one in the WT1 lane, it can be seen to resolve at a different molecular weight. The 12CA5 lane shows this difference more clearly because it has both bands present. It has the higher band, running at the same level as the band in the WCE and representing HA-NAL, and it has the lower band running at the same level as that in the WT1 track and representing the heavy chain of the antibody. The negative control IgG also has slight band, and again this would represent the heavy chain of the antibody and not HA-NAL, and this runs at the

same level as the WT1 immunoprecipitate band. Thus this experiment did not show that WT1 and HA-NAL interacted.

The experiment was repeated, but again there was no HA-NAL band seen in the WT1 immunoprecipitate. However, it may not be too surprising that the experiment cannot detect WT1 immunoprecipitating HA-NAL. Apart from the fact that perhaps the interaction between NAL and WT1 could be transitory, this experiment was carried out following transient transfections. Therefore not all cells would have the HA-NAL protein, and for those cells that do, perhaps not all of the HA-NAL would be involved in binding to WT1, and so the levels of HA-NAL protein to be detected in the WT1 immunoprecipitate could be fairly low.

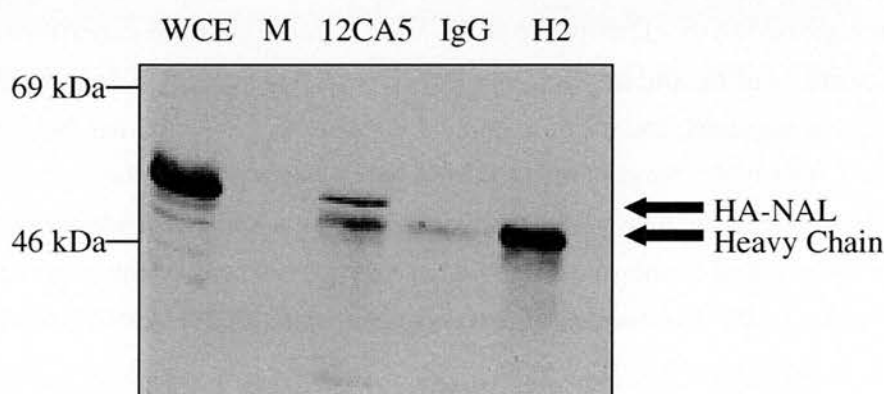


Figure 5.8: Co-immunoprecipitations of transiently transfected HA-NAL. The immunoblot was blotted with 12CA5, an antibody which would recognise HA-NAL which had been transiently transfected into M15 cells. The whole cell extract (WCE) lane shows the expression of HA-NAL in the transfected M15 cells. The 12CA5 lane shows that the immunoprecipitation conditions were correct as this antibody did immunoprecipitate NAL. The H2 lane shows the immunoprecipitation carried with the H2 antibody directed against WT1. If NAL and WT1 interacted, there should be a band representing NAL, as NAL would have been part of the WT1-complex. Although there is a band present in the lane, it is likely to be a non-specific band due to the detection of the heavy chain of uncrosslinked antibody. This band can be seen in the negative control, IgG, and also it is the lower band in the 12CA5 track. (M=markers)

As it was hard to detect an interaction between the two proteins in these transient transfection experiments, it seemed logical to try to detect the association in cells that endogenously express both proteins. An antibody to NAL would therefore have to be made, which in the long run would prove beneficial as it could not only be used for these *in vivo* experiments, but would also provide a resource for future experiments.

5.3 Producing A Polyclonal Antibody Against NAL

5.3.1 Production Of Antibody

An antibody against NAL was required, and it was decided to raise polyclonal antisera in rabbits. The rabbits were to be injected with the NAL protein such that the immune response would produce antibodies which recognise NAL. Originally NAL was to be made by bacterially expressing the GST-NAL clone as described in section 5.1.3. Then the GST-NAL fusion protein was to be cleaved to remove the GST portion, and this was to be done by taking advantage of the thrombin cleavage site found after the GST moiety. However, only a low yield of NAL was recovered after thrombin cleavage (data not shown) and so it was decided to inject GST-NAL into the rabbits instead. The immunisation of rabbits and the isolation of the serum was carried out by the Scottish Antibody Production Unit (SAPU) as detailed in section 2.7.1. The rabbit antiserum was then affinity purified to get a good quality antibody to NAL as described in section 2.7.2.

5.3.2 Testing The NAL Antibody

The affinity purified NAL antibody, called 474, had to be tested to ensure it recognised NAL, and to determine the concentration it should be used at for subsequent experiments. M15 cells were again transiently transfected with HA-NAL and after 48 hours of expression, the cell lysate was run on an 8% SDS PAGE. Cell lysates from untransfected M15 cells were also run, as they express endogenous NAL.

Three immunoblots, each containing cell extracts of both transfected and untransfected M15 cells, were probed with different dilutions of the 474 antibody against NAL. As shown in figure 5.9, the 474 antibody can recognise NAL in both its endogenous form and as HA-NAL. The band produced for untransfected M15 cells is very clean with no other background. This probably reflects the purity of GST-NAL initially injected into the rabbits, leading to a very specific immune response. The transfected M15 cells produced a slightly larger sized band than for

untransfected cells because the transfected cells had the HA epitope tagged onto NAL. The other bands appearing in the track are probably breakdown products of the excess HA-NAL. As seen from the dilution assays done, the best concentration to use 474 to detect endogenous NAL is between the 1 in 200 or the 1 in 500 dilution.

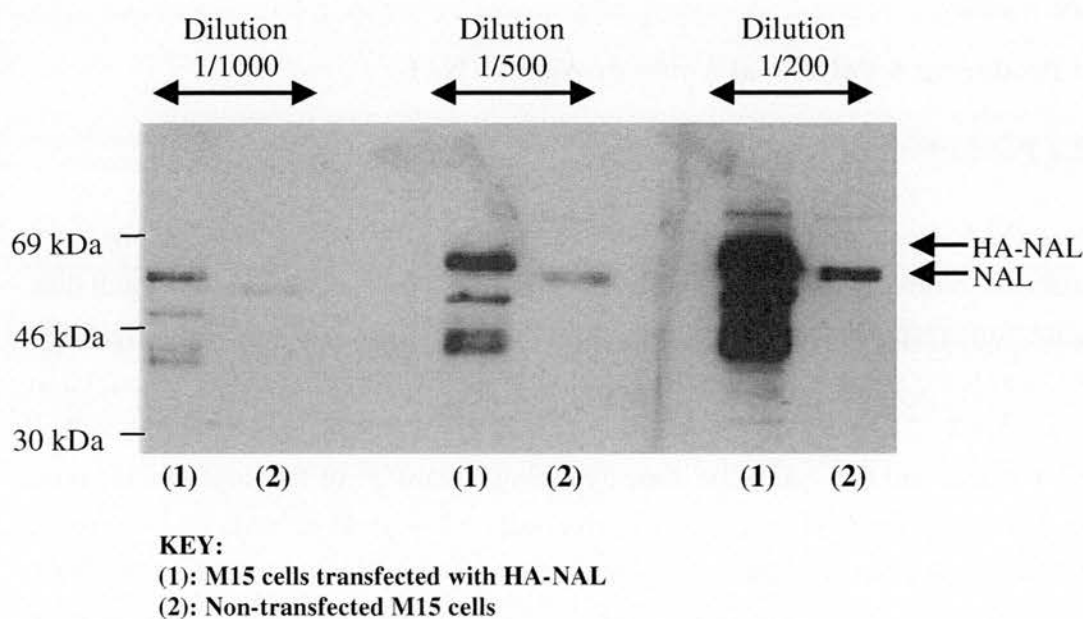


Figure 5.9: Immunoblot showing that the NAL antibody can detect both endogenous and transfected NAL in M15 cells. 3 immunoblots, each containing whole cell extracts from either M15 cells transfected with HA-NAL, lane (1), or non-transfected M15 cells, lane (2), were probed with the NAL antibody 474 used at different dilutions. All 3 blots could detect the endogenous and the transfected NAL but the best dilution to use the antibody for detection on immunoblots is between 1/500 to 1/200.

The 474 antibody recognises endogenous NAL which is resolved at a molecular weight of about 55kDa, as shown in figure 5.9. This molecular weight is the same as that produced when the protein is made *in vitro* from the NAL clone that was isolated from the yeast two hybrid. This can be seen in figure 5.9, where the HA-NAL fusion protein, which is only just larger than the NAL clone due to the HA tag, runs slightly higher than the endogenous NAL protein in M15 cells. This therefore indicates that the NAL clone isolated is very nearly full length, as discussed in section 4.5.

5.4 *In Vivo* Binding Assays using Endogenously Expressed NAL

Now that there was an antibody that recognised endogenous NAL, the *in vivo* experiments that would assess whether NAL binds to WT1 under physiologically relevant conditions, and not just in yeast or *in vitro*, could be done. Co-immunoprecipitations were carried out on M15 cells essentially as detailed in section 5.2.3. This time no transient transfections were carried out, and the antibodies used to carry out the immunoprecipitations were 474 to test whether this antibody could immunoprecipitate NAL, and if it could it would act as a positive control; IgG as a negative control; and WT1 to assess whether binding between NAL and WT1 was occurring.

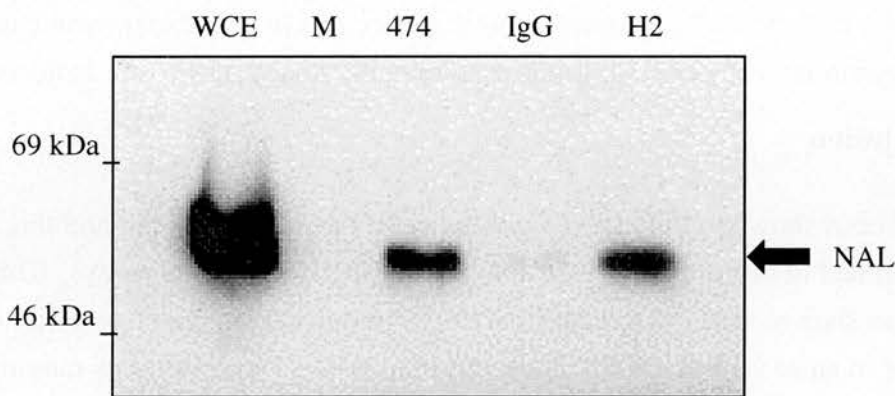


Figure 5.10: WT1 co-immunoprecipitates NAL. The immunoblot was probed with the NAL antibody 474. The whole cell extract (WCE) lane shows that the expression of NAL can be detected by 474. The 474 track represents an immunoprecipitation carried out with 474, and the band seen is the immunoprecipitated NAL. Thus the positive control showed that the immunoprecipitation reactions were working. The IgG track represents an immunoprecipitation carried out with the non specific mouse IgG antibody which served as a negative control. Finally the H2 track represents an immunoprecipitation carried out with the anti-WT1 antibody H2. The band present in this track is NAL, which means that it was co-precipitated with WT1 and so the two proteins must interact.

Figure 5.10 shows that now an association between NAL and WT1 could be detected. The immunoblot was carried out with 474 antibody, to detect NAL. The lane containing the whole cell extract (WCE) shows that the endogenous NAL could be detected. The immunoprecipitation carried out with 474 also shows a band for NAL, which demonstrates that the antibody can not only be used for immunoblots,

but also for immunoprecipitations. The immunoprecipitation with 474 also acts as a positive control showing that the conditions for the immunoprecipitations were good. The track which immunoprecipitated WT1 with the H2 antibody also shows a band representing NAL, whereas the negative control, IgG, doesn't. This shows that WT1 and NAL must have interacted for it to be present in the WT1 immunoprecipitate track. An identical blot (not shown) had also been probed with the secondary antibody only, without having first blotted with 474, to ensure that the bands produced were not non-specific. This blot did not produce any signal at all, showing that the bands seen in figure 5.10 were representing NAL. This *in vivo* experiment therefore shows that WT1 and NAL can interact under physiologically relevant conditions. Unfortunately there are no antibodies that reliably distinguish between the different WT1 isoforms, and therefore this *in vivo* experiment cannot be used to determine whether NAL binds preferentially to any of the four isoforms.

5.5 Conclusion

NAL had been shown to bind to WT1 in the yeast two hybrid system and this interaction had to be confirmed by using both *in vitro* and *in vivo* assays. This chapter has shown, after overcoming several problems due to the insolubility of WT1, that in an *in vitro* situation, NAL can bind to WT1 in conditions ranging from low to high stringency, suggesting that the interaction between the two proteins must be fairly strong. The *in vitro* data agreed with the results produced in the yeast two hybrid system, in that exons 1 and 2 of WT1 are not needed for the interaction between NAL and WT1 to occur, as the deletion construct Nco5' did not bind to NAL, whereas the deletion constructs of Nco3' did. The *in vitro* experiments did not show any apparent difference in binding affinities between the different WT1 isoforms.

The interactions found in the yeast two hybrid system and *in vitro* may not reflect what is going on at the cellular level because the two proteins are overexpressed in these assays. The binding assays therefore had to be carried out *in vivo*, to ensure that the interaction was occurring at the cellular level. Initially these experiments could not be carried out using endogenous NAL because there were no antibodies available to detect this protein. For this reason transient transfections of HA epitope-tagged NAL had to be carried out. Antibodies to the HA tag were then used to assess whether HA-NAL was found in WT1 complexes. Unfortunately the experiment did not show any HA-NAL present in WT1 complexes. This was probably because under transient transfection conditions, relatively few cells would

be expressing the HA-NAL and so even if HA-NAL was found in WT1 complexes in these cells, the levels may have been below detection.

Since the transient transfections were not adequate, an antibody to NAL had to be produced to ensure that the *in vivo* assays, with endogenous NAL, could be carried out. A polyclonal antibody was produced, 474, and it proved to be very specific to NAL, giving very clean Western blots. The antibody made was also able to immunoprecipitate NAL. This antibody was therefore used for the *in vivo* experiments, which now showed that endogenous NAL could be found in complexes co-immunoprecipitated with WT1. These experiments all show that WT1 can bind to NAL and that the interaction is genuine, occurring under *in vivo* conditions.

Chapter 6

Expression And Cellular Localisation Studies

Chapter 6: Expression and Cellular Localisation Studies

NAL was isolated due to its ability to bind to WT1 in the yeast two hybrid system. The association between the two proteins was also shown to occur *in vivo* and thus it was important to find out more about the novel protein to understand its role at the molecular level and why it was binding to WT1. At this stage not much was known about NAL as its protein sequence provided no clues towards its possible function due to a lack of known motifs within it. The mouse homologue of NAL was isolated, and it was seen that the sequence was highly conserved between mouse and human, suggesting that the structure of the protein is likely to be important for function. As discussed in chapter 4, mapping studies on both human and mouse NAL were carried out and it was found that the human NAL mapped to human chromosome 6q26-27, and the mouse Nal mapped to chromosome 17, very close to the centromere. Although the data are interesting, especially as the human NAL could be mapping to a site where there is thought to be a tumour suppressor gene, it does not provide any clues as to its possible function and why it binds to WT1. Studies were therefore carried out to try to further characterise this novel protein, and these included expression studies to determine whether the gene is expressed only in certain tissues, as WT1 is, or whether it is ubiquitously expressed and therefore perhaps carries out a housekeeping function in every cell. Localisation studies within the cell were also carried out to determine whether the protein is nuclear or cytoplasmic and from this knowledge, further experiments could then be devised to start to decipher its role.

6.1 Determining the expression pattern using RT-PCR

RT-PCR was used to determine which tissues expressed NAL. The experiment was carried out on RNA that had been isolated from tissues of an adult female mouse as is discussed in section 2.4.2. The primer pair used to carry out the amplification were primers L342, starting at base 416 of NAL, and M718 starting at base 1113 - the primer sequences used can be found in the appendix.

Figure 6.1 shows that NAL is expressed in all tissues that were studied and these included: brain; thymus; heart; lung; liver; spleen; kidney; and muscle. The results show that the main RT-PCR product was about 700 bases as expected but some tissues have other smaller bands present too. These bands could represent alternative transcripts, which vary between different tissues as the lung and the thymus both have a band at about 550 bases, and although very faint, traces of the same band appear in the brain, spleen and kidney samples whereas the heart and

possibly the lung on the other hand, seem to have a product in the 300 base range. The idea that there could be alternative transcripts of NAL was already discussed in section 4.2, as an EST was found in the database which seems to be an alternative transcript of NAL. Whether these other bands really represent alternative transcripts is not known as the products could have been derived non-specifically, or they may also represent transcripts from a closely related gene. As an approach to obtain more specific PCR products, nested PCR could have been carried out. Alternatively, to determine unequivocally whether the smaller amplified PCR products are alternative transcripts of NAL, the amplified bands should have been cloned and sequenced.

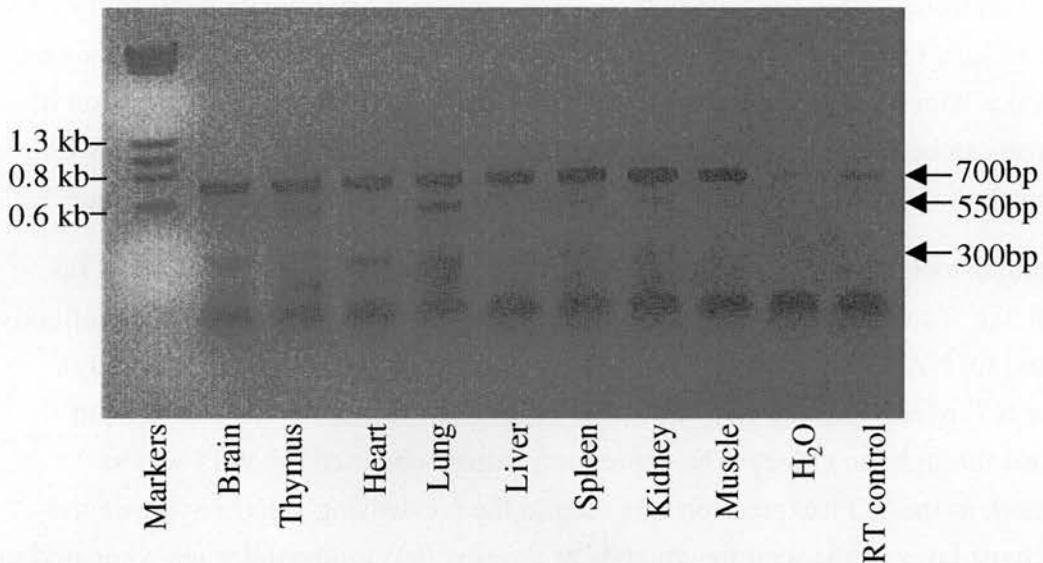


Figure 6.1: Expression studies for NAL using RT-PCR. Tissues from an adult mouse were used to test the expression pattern of NAL. The figure shows the products amplified using primers L342 and M718. The expected 700 base product is amplified in all tissues and not in the negative controls. The RT control lane is a negative control, where RNA from the kidney had been used but no reverse transcriptase had been added so that no cDNA could have been produced. This control therefore shows that the PCR products are not being amplified from contaminating genomic DNA. Apart from the main 700bp band, other smaller bands, possibly representing alternative transcripts can be seen for the thymus and lung (550 bp) as well as for the heart (300 bp). The nature of these smaller products was not looked at and so they could equally be non-specific bands which are being amplified.

These RT-PCR experiments show that NAL is expressed in a wide range of tissues of the adult mouse. This is unlike WT1, which in the adult seems to be expressed only at low levels in the spleen; heart; gonad; kidney; tongue and eye (Armstrong et

al., 1993; Buckler et al., 1991). NAL is therefore not tissue specific and so its expression is not confined to organs which express WT1. This suggests that NAL and WT1 could be interacting in organs where both proteins are expressed but because NAL is also expressed in organs where WT1 is not found, it is likely to also function independently of WT1.

6.2 Immunofluorescence on sections through the kidney and testes

The RT-PCR experiment shows that NAL is expressed in the adult kidney. It was of interest to examine whether NAL is also expressed in the developing kidney and if so whether, like WT1, it is confined to the developing nephros - in other words, is NAL localising within the kidney just to areas where it can interact with WT1? To examine this, kidneys from mice at E=17.5d were used, as at this age the kidney has reached a stage in development where all the structures leading to the formation of a nephron can be seen. Section 1.7.1 described the expression pattern of WT1 within the developing kidney, and the structures expressing WT1 are depicted in figure 1.2 .

Immunohistochemistry was performed on cryosections of kidneys at E=17.5d (as detailed in section 2.9.2). The kidney sections were stained for WT1, using antibody H2, and for NAL, using antibody 474 (see chapter 5 for details of this antibody). Figure 6.2, panels (A) and (B), shows the expression seen for NAL and WT1 in sections through the kidney. The expression pattern obtained for WT1 was as expected, as the WT1 expression was seen in the condensing mesenchyme on the outer most layer of the section, which is at an early developmental stage compared to the inner layers of the kidney. As the glomerulus develops further, the expression of WT1 can be seen in the Comma shaped bodies, followed by the S-shaped bodies and finally being confined to the podocytes of the glomerulus (see figure 6.2 C). In contrast to the confined expression pattern of WT1, NAL is shown to be expressed throughout the whole kidney, including those areas where WT1 is expressed (seen as a yellow output in the merged panel in figure 6.2 A and B). The expression of NAL is therefore not confined to the structures within the kidney that express WT1.

The expression of NAL was also looked at in the testes as shown in figure 6.2 D. The testes were dissected out from the E=17.5d mice and immunohistochemistry was performed using the same antibodies as were used for the kidneys. The expression pattern for WT1 was again as expected, with WT1 being expressed mainly in the Sertoli cells, and not in the germ cells. The expression of NAL again is ubiquitous within the testes. The expression overlapped with WT1 within the Sertoli cells, but NAL was also expressed in the germ cells and throughout the rest of the testes.

These immunofluorescence studies therefore show that NAL is expressed within the kidney and testes, but rather than having a restricted expression pattern within the developing nephron and Sertoli cells like WT1, NAL is expressed throughout the whole of these organs.

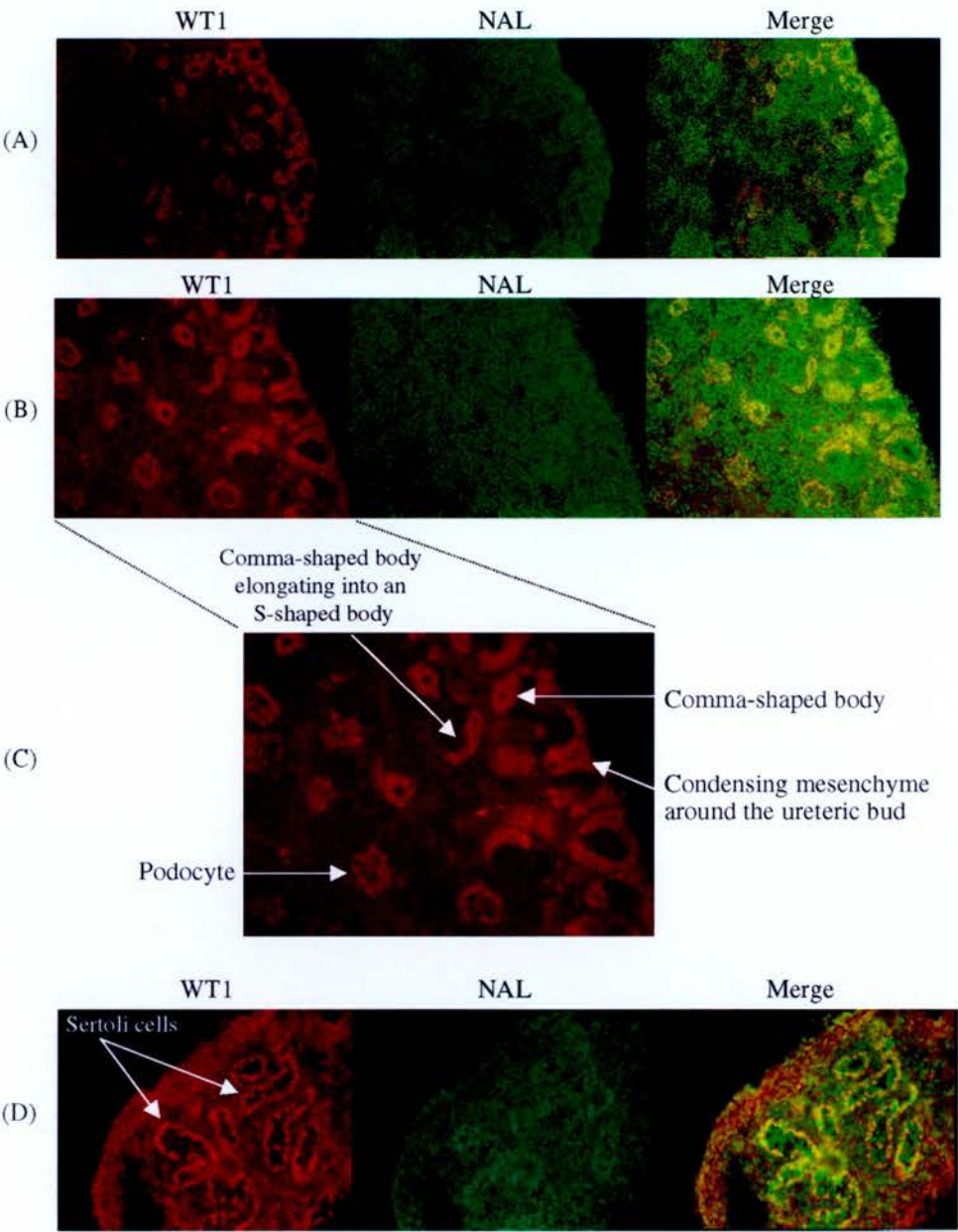


Figure 6.2: Immunofluorescence on sections through the kidney and testis. Panels (A) and (B) show sections through the kidney at lower and higher magnification respectively. The immunofluorescence produced using antibodies against WT1 (antibody H2; red output) and NAL (antibody 474; green output) is shown, as is their merged output to see where they co-localise. Panel (C) highlights those areas within the developing nephros where WT1 is expressed, as is described in figure 1.2. Panel (D) shows the expression pattern of WT1 and NAL, together with their merged output, within a section through the testis.

6.3 *In situ* hybridisation studies

6.3.1 Whole mount *in situ* hybridisation

The RT-PCR experiments provided some idea as to the type of tissues where NAL is expressed, but for a more detailed study of both temporal and spatial expression, whole mount *in situ* hybridisation experiments were carried out. Whole mount *in situ* hybridisation is a technique used to detect the presence of the RNA of the gene of interest in the context of whole embryos. It uses an antisense RNA probe, which in this case was transcribed from the mouse *Nal* clone isolated (see section 4.4), that is labelled with digoxigenin (DIG) as described in section 2.4.3. The embryos are then hybridised with this antisense *Nal* probe, as discussed fully in section 2.5.1. The antisense *Nal* probe will hybridise with endogenous *Nal* RNA wherever it is expressed. Any probe which has not hybridised with endogenous *Nal* RNA, will be washed away in the subsequent steps of the protocol. The hybridised probe can be detected by enzyme-linked immunoassays using an anti-digoxigenin alkaline phosphatase conjugate (anti-DIG-AP). A subsequent enzyme-catalysed colour reaction with 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) and nitroblue tetrazolium salt (NBT) produces an insoluble blue precipitate, thus enabling detection of the hybrid molecules. As a control, this experiment can be carried out with a sense RNA *Nal* probe. This sense probe cannot hybridise to the endogenous *Nal* RNA in the embryo, and so is removed during the washing steps. Thus any non specific signal detected will also show up on this negative control.

Mouse embryos between embryonic ages E= 9.5d-12.5d were used for these studies, and both sense and antisense *Nal* RNA probes were used for all ages except the E=11.5d where only the antisense probe was used. The results produced are shown in figure 6.3, in which a whole mount *in situ* for an antisense WT1 probe at E=9d is also shown (taken from Jacob Hecksher-Sørensen and Adrian Moore) . Although only one embryo for each of the probes is shown at each stage, a total of 3 embryos were stained for each probe used at the different embryonal stages to ensure that the signal obtained was real. What is immediately obvious from these results is that *Nal* seems to be expressed throughout the embryo at all the ages tested. As the sense control was used, it can be seen that the expression observed for *Nal* is real and is not just background staining of the embryo (apart from a few areas in the E=10.5d embryo where there is a little 'trapping' of the probe in the brain and neck). Apart from being ubiquitously expressed, there are a few other areas where the expression levels seem to be slightly higher, as is the case for the limbs; at the tip of the tail end; and in the E=9.5d in the nasal process and mandibular component of the first

branchial arch, which in the older embryos the tissues derived from these components also have a higher expression level.

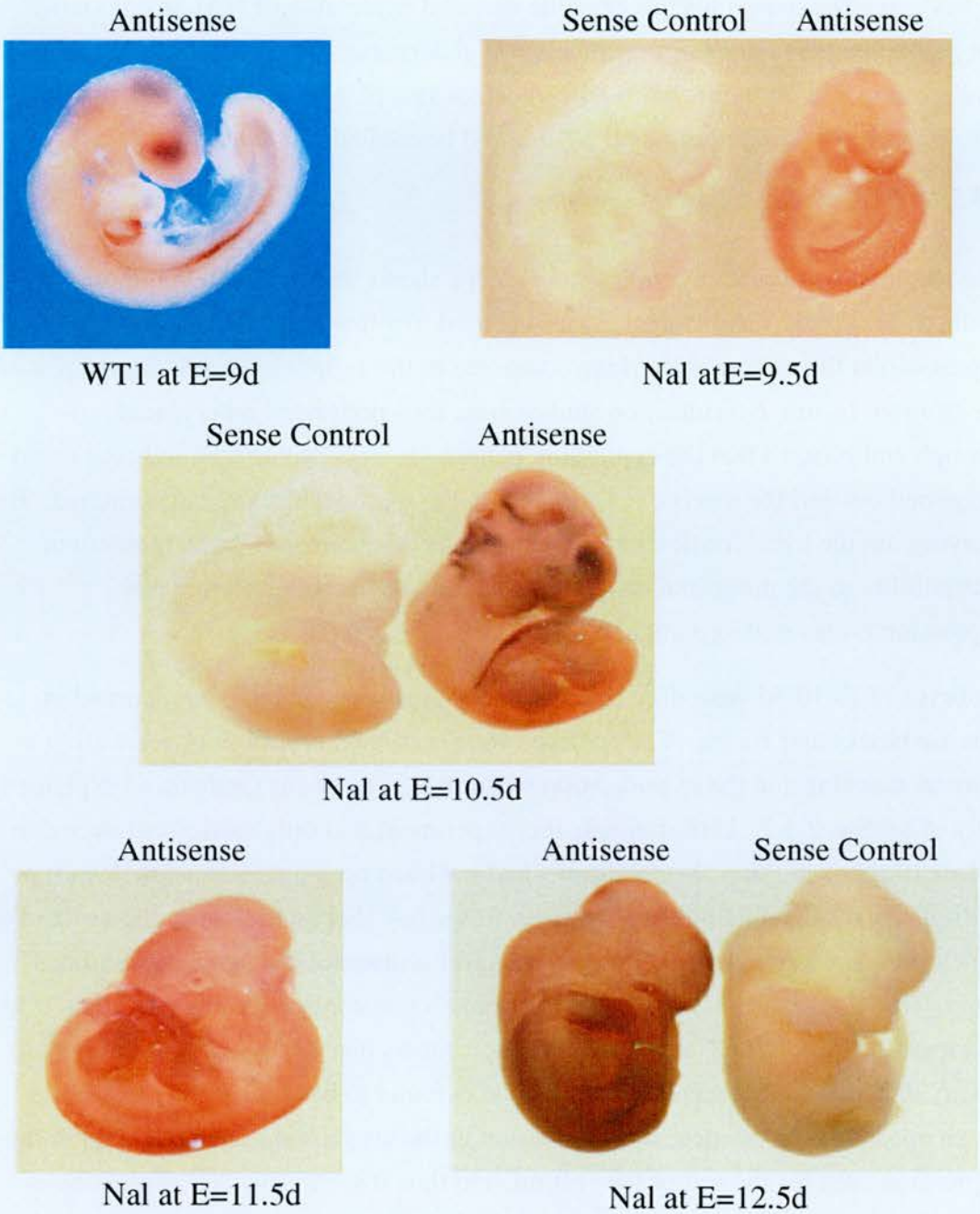


Figure 6.3: Whole Mount *In Situ* Hybridisation. These whole mount *in situ* hybridisations show the RNA expression pattern of WT1 and NAL. The WT1 expression is only shown for E=9d (taken from Moore et al 1998). Embryonic ages between 9.5d-12.5d were used to see if there is any variation in the expression pattern at the different developmental stages.

Although these areas could simply be sites to which the probe has easier access and therefore produce a stronger signal, they could be genuine sites of higher NAL expression. The reason for this possible elevated expression of NAL is not known, but perhaps these regions of the embryo are undergoing a lot of cellular proliferation, although this has not been demonstrated, and thus a possible link between cellular proliferation and NAL could be studied in the future.

6.3.2 *In situ* hybridisation on sections

The whole mount *in situ* hybridisation of WT1 shows that the RNA is found mainly in the mesonephric ridge (figure 6.3). It is hard to tell whether Nal shows a stronger expression at the mesonephric ridge compared to the expression levels for the rest of the embryo. *In situ* hybridisation studies were thus performed on cryosections through embryos, so that the expression pattern on cross-sections of embryos could be carried out and the levels of expression on the mesonephric ridge determined. By carrying out the hybridisation on sections means that there will be no problem of accessibility to the probe and so it will be easier to determine whether Nal expression is elevated in particular areas.

Embryos of E=10.5d were dissected out and immediately fixed, then mounted in agarose blocks and frozen. The embryos were sectioned (10µm thickness) using a cryostat machine and the hybridisation was carried out by Liz Graham, as explained fully in section 2.5.2. Unfortunately this experiment was only carried out once due to lack of time, and thus the conditions had not been optimised. Nevertheless, the preliminary results obtained are shown in figure 6.4. Sections close to the tail-end of the embryo are shown for the three probes: Wt1 antisense; Nal antisense; and Nal sense. The Wt1 section shows that Wt1 expression is confined to the urogenital ridge and mesothelium. The Nal antisense section shows that the Nal expression pattern is very different to that seen for Wt1, as Nal is found to be ubiquitously expressed. There appears to be no increased expression in the urogenital ridge compared to the expression seen for the rest of the section, and thus it seems that Nal expression is not upregulated in areas which express Wt1. On these sections there does not appear to be any stronger expression of Nal in the limbs and so this suggests that the stronger expression levels in the limbs and in the posterior end of the tail seen in the whole mount *in situ*, may have been due to the probe having better accessibility to these areas.

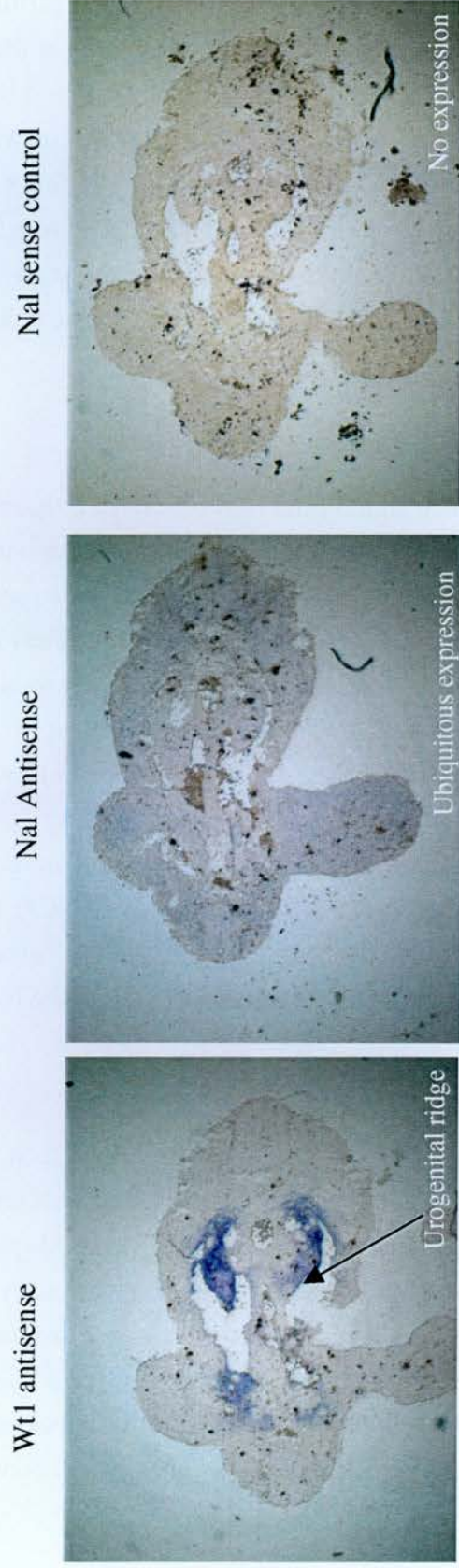


Figure 6.4: *In situ* hybridisation on mouse E=10.5d cryosections. This figure shows the differences in expression between Wt1 and Nal. Wt1 expression is confined to urogenital ridge and the mesothelium, whereas the Nal expression, as demonstrated by the antisense probe, is throughout the whole section. The expression of Nal seems to be weaker than that of Wt1, but by comparing it to the Nal sense control, there is a definite expression seen. Contrary to what was seen in the whole mount *in situ* hybridisation, on these sections there seems to be no areas where the Nal gene is expressed to a greater extent. However it must be stressed that these were preliminary experiments, and the staining conditions may not have been optimised for Nal.

6.4 NAL is a nuclear protein

The cellular localisation of NAL was looked at to determine in which compartments NAL was mainly present. It was important to know whether at least some of the protein was nuclear, as if it is to interact with WT1, then both proteins should be in the same cellular localisation. Immunocytochemistry was thus carried out on M15 cells, as these cells express both NAL and WT1 - which was required for direct comparisons between the endogenous localisation of WT1 and NAL. Figure 6.5 (panel A) shows that NAL, like WT1, is a totally nuclear protein. This was an encouraging finding as it shows that WT1 and NAL are present in the same cellular localisation and thus are able to interact.

6.5 NAL co-localises with splicing factors

Closer inspection of the immunofluorescence signal produced for NAL in figure 6.5 A, shows that the pattern within the nucleus appears to be speckled together with an overall diffuse stain which excludes the nucleolus. This was the pattern also seen for WT1 and which had previously been described by Larsson *et al.* (1995). They then went on to show that the speckle structures in which WT1 was found corresponded to sites which co-localised with splicing factors. It was therefore of interest to determine whether the speckles in which NAL are found in are the same as those speckles seen when the nucleus is stained for splicing factors, and thus immunocytochemistry was carried out on M15 cells with the monoclonal antibody 3C5. This antibody detects proteins in interchromatin granule clusters (IGCs), the structures in which splicing factors are stored (Misteli and Spector, 1998; Turner and Franchi, 1987). The immunocytochemistry was also performed with the NAL antibody, 474, to see if the speckles seen for NAL co-localise with those of splicing factors.

Figure 6.5 B shows the immunocytochemistry results produced for 3C5 (green, FITC signal) and NAL (Texas Red signal). The 3C5 panel shows large foci which are the storage sites of splice factors, together with some diffuse staining. The NAL panel shows that NAL is also present in the same speckled structures, which are seen as a yellow output in the merged panel. The speckled staining of NAL is not as intense as seen for 3C5, and this could perhaps reflect the fact that under these conditions, NAL may be preferentially localising to the nucleoplasm instead of the speckles. Nevertheless the speckles in which some of NAL are the same ones which contain splicing factors.

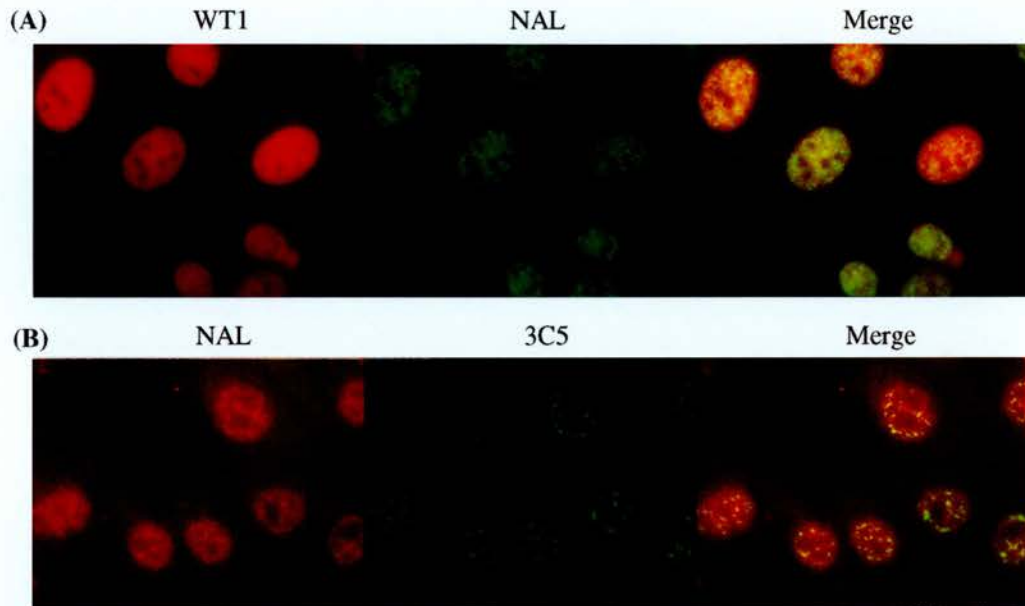
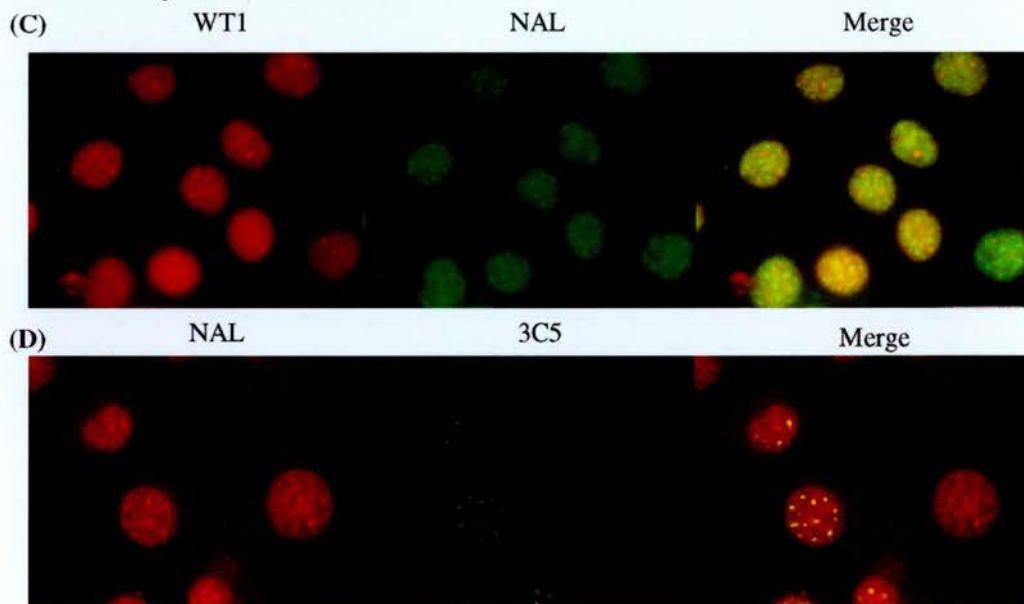
I: No Treatment**II: Actinomycin D Treatment**

Figure 6.5: Immunofluorescence on M15 cells. Panel (A) shows that both WT1 (antibody H2; red output) and NAL (antibody 474; green output) are nuclear. The staining pattern for both WT1 and NAL seems to be speckled together with an overall diffuse stain. Panel (B) shows that the speckles seen for NAL (red output) co-localise with those of 3C5 (green output) which stains interchromatin granule clusters, where splicing factors are found. After actinomycin D treatment, NAL behaves differently to WT1 as shown in panel (C). WT1 re-localises to a few, small foci around the remnants of the nucleoli (red output), where as NAL goes into several larger foci (green output). Panel (D) shows that the large foci where NAL is found in after actinomycin D treatment (red output), co-localise with the foci stained by 3C5 which contain the bulk of the splicing factors.

6.6 NAL redistributes into large foci after actinomycin D treatment

The interchromatin granule clusters are thought to be sites where splicing factors are stored and/or assembled until they are required for processing pre-mRNAs. At this point they are recruited out of the speckles to the sites of active transcription at the perichromatin fibrils (Misteli et al., 1997; Spector, 1996). The experiments described in these two papers show that the speckles are very dynamic, as splicing factors shuttle between the speckles and the site of active transcription. However when transcription is inhibited, the splicing factors are not required and so remain in the IGC, causing these structures to become larger and more uniform in shape. Not all splicing factors return to the IGC when transcription is inhibited, as changes in localisation are to a large extent splice factor dependent (Spector et al., 1991). For instance U170K and U2AF65 are examples of splicing factors which instead of being found in the large foci after inhibition of transcription, are present in small foci around the remnants of the nucleolus (Carmo-Fonseca et al., 1992). The same was seen to be true for WT1, as when M15 cells were treated with the transcription inhibitor actinomycin D, WT1 localised to small foci which co-localised with U170K and U2AF65, and not with the other snRNPs (Larsson et al., 1995).

The nuclear localisation of NAL was therefore looked at after M15 cells had been treated with actinomycin D, to see if the treatment affected the nuclear distribution of NAL. It was of interest to determine whether, if NAL did redistribute, it behaved like WT1 and localised to small foci around the nucleolus, or whether it was found in several large foci where most of the other snRNPs are present. Figures 6.5 C and D, clearly show that when M15 cells are treated with actinomycin D, NAL redistributes into large foci. Panel 6.5 C shows that WT1 is present in small foci found around the remnants of the nucleolus, as seen previously by Larsson *et al.*, 1995. These foci are different to the large, abundant foci seen for NAL, and as the merged output shows, the foci in which WT1 are found in do not co-localise with the foci in which NAL are in. Panel 6.5 D on the other hand shows that the large foci in which NAL are found in do co-localise with the large foci stained by 3C5. The staining seen for NAL does vary to that seen for 3C5 in that even after actinomycin D treatment, NAL does continue to have a substantial amount of nucleoplasm staining, whereas for 3C5 there is hardly any nucleoplasm staining at all, as the splicing factors are localised almost exclusively to the speckles.

These experiments have therefore shown that NAL is a nuclear protein which is found to be present in speckles as well as being distributed throughout the nucleoplasm. This nuclear pattern is what has been seen for WT1 (this study and

(Larsson et al., 1995)). However, when the cells are treated with the transcription inhibitor, actinomycin D, WT1 and NAL behave differently. WT1 becomes localised to small foci around the remnants of the nucleolus, as is seen for U170K and U2AF65, whereas NAL goes into large foci which co-localise with the other snRNPs.

6.6 Conclusions

This chapter has investigated the expression pattern of NAL as well looking at the cellular distribution of the protein. The expression of NAL was looked at and compared to the expression pattern previously documented for WT1. Both RT-PCR experiments and *in situ* hybridisation techniques have shown that NAL is ubiquitously expressed, therefore being expressed in tissues where WT1 is not found. The immunohistochemistry carried out on kidneys and testes showed that in these organs, where both NAL and WT1 are expressed, the expression pattern of NAL is not identical to that of WT1. While WT1 is confined to certain structures within these organs, NAL is expressed throughout both organs. Thus NAL and WT1 seem to be co-expressed but only in certain tissues and within these, only in particular structures. Due to its ubiquitous expression, NAL may have a housekeeping role, and in cells which also express WT1, the function of NAL may be altered due to its interaction with WT1, and/or NAL may in turn affect the function of WT1.

Using immunocytochemistry NAL was shown to be a nuclear protein and within the nucleus it was shown to have both a diffuse distribution throughout the nucleoplasm, excluding the nucleolus, as well as a speckled localisation. This is what had previously been reported for WT1 (Larsson et al., 1995). This was a very encouraging finding as it means that WT1 and NAL are in the same cellular compartments and are therefore able to interact. Moreover it was interesting to find that NAL also had a speckled pattern within the diffuse staining in the nucleoplasm, suggesting that perhaps it may interact with all the different WT1 isoforms. However, WT1 and NAL differ in their response to treatment with the transcription inhibitor actinomycin D. When transcription is inhibited, WT1 localises into small foci which are found around the remnants of the nucleoli, just like U2AF65, whereas NAL goes into larger speckles which were shown to be the interchromatin granule clusters, where the bulk of the splicing factors are stored. Although NAL is found in these large speckles, it does not confirm that it itself is a splicing factor. The function of these speckles is not yet really understood, but it is thought that these sites are where splicing factors are stored and perhaps assembled into spliceosomes

before they are recruited to the sites where the splicing reactions occur (Misteli et al., 1997). It is thought that apart from the components of the spliceosome, there may be other proteins present within the speckles which could have nothing to do with splicing *per se*, and perhaps NAL could be one of these. It has been noted that many of the splicing factors that are recruited into the speckles have an arginine/serine (RS) rich motif (Li and Bingham, 1991) or an RNA binding motif together with a certain combination of other motifs (Caceres et al., 1997; Gama-Carvalho et al., 1997). The exact combination of motifs required to send a protein into the speckles is not really understood and thus it is of interest to note that NAL does not have an RS motif or an RNA-binding motif but still localises to the speckles.

As the immunocytochemistry of NAL overlapped with that of 3C5, and therefore with splicing factors, the possibility that NAL could play a role in splicing cannot be ruled out. The role of WT1 in splicing is not yet known, and indeed the question of whether WT1 is involved in splicing at all is still very controversial. However, if NAL and WT1 do play a role in splicing, and since they have been shown to interact *in vivo*, it is perhaps surprising that WT1 and NAL did not co-localise after actinomycin D treatment. This need not be a problem when trying to determine whether the interaction between NAL and WT1 is significant in a splicing context because not much is known about the assembly of the spliceosome. For instance it is not certain whether the spliceosome assembles in the speckles before being recruited to the site where it is required, or whether splice factors are recruited individually from the speckles as they are needed. Therefore if WT1 and NAL do have a role in splicing, they do not necessarily have to be interacting within the speckles, but only at the site where splicing occurs.

What is known about splicing is that the system requires several snRNPs and accessory proteins, whose interaction leads to the removal of introns by sequentially binding to the pre-mRNA. It is first the U1 snRNP which binds to the 5' splice site of the pre-mRNA. The accessory factor U2AF65 then promotes the binding of the U2 snRNP to the pre-mRNAs 3' branch point (Zamore et al., 1992). Once U2 is bound to the branch point, which occurs only after U1 has bound to the 5' splice site, it commits the pre-mRNA to splicing. The interaction between U1 and U2 brings the two splice sites together. The trimeric complex of U5 and U4/U6 then completes the assembly of the spliceosome and splicing can begin. The splicing reaction is triggered by the release of U4 from the complex, which probably acts to sequester U6 until it is required. The release of U4 now allows the RNA component of U6 to hybridise to part of the RNA within U2, the other part being hybridised to the branch point of the pre-mRNA. The U5 recognises the 3' splice site although it is not yet

known if it does so directly. The splicing reaction then occurs by two transesterification events, leading to the joining of the two exons through an intermediate lariat structure.

Davies *et al.* speculate that perhaps WT1 influences 3' splice site selection by binding to a 3' branch point through its RRM and presenting it to U2AF65 by directly interacting with it (Davies *et al.*, 1998). As NAL interacts with WT1, it could have a role at the branch point too, perhaps stabilising the complex between WT1, U2AF65 and U2. However, it has also been suggested that WT1 may act at several different stages during splicing (Davies *et al.*, 1998) and so NAL could be interacting with WT1 at any of these points. There are several as yet unidentified non-snRNP splicing factors within the spliceosome and therefore NAL could be one of these, acting at any of the stages during splicing. Because NAL is ubiquitously expressed, it is plausible that it is part of the spliceosome and that perhaps WT1 binds to NAL to somehow affect the outcome of the overall splicing event.

The immunocytochemistry showed that NAL, apart from being present in speckles, also had a diffuse staining pattern. Splicing factors also have a slight diffuse pattern as well as the speckled pattern, as the factors do get recruited out of the speckles to the site of splicing. However, after treatment with actinomycin D, the splicing factors tend to return to the speckles, thus not only enlarging these structures but also removing most of the diffuse staining. On the other hand, even after treatment with actinomycin D, NAL still showed a diffuse staining pattern as well as the enlarged speckles suggesting that a lot of the protein is not involved with the speckles. This could perhaps suggest that NAL could have two roles, such as splicing and transcription as has been postulated for WT1. The different roles of NAL could arise due to its potential alternative transcripts, which were mentioned earlier and in chapter 4. The proteins derived from these different transcripts would possibly be detected by the NAL antibody in the immunocytochemistry experiments because it was raised against the whole protein as opposed to specific peptides.

The whole mount *in situ* hybridisation experiment carried out for NAL suggested that perhaps NAL is expressed higher in areas which may be proliferating at an increased rate. As described in section 1.6.4, WT1 plays a role in cell cycle regulation. In the kidney WT1 is thought to inhibit cellular proliferation and to induce cells to differentiate into epithelial cells. It may be that the role of NAL could be associated with the cell cycle, and that perhaps WT1 interacts with NAL to bring about the desired effect. Thus it would be of interest to determine whether NAL does have a role in the cell cycle.

Unfortunately due to lack of time, no real functional assays could be done, and so the role of NAL at the molecular level is not yet clear. However, its distribution within the nucleus suggests that perhaps NAL could be playing one than one role, and these may be brought about by alternative transcripts. This study has raised several speculative functions of NAL but its true role and the consequences of its interaction with WT1 remains to be investigated.

Chapter 7

Summary and Discussion

Chapter 7: Summary and discussion

This project set out to identify WT1 interacting proteins, in the hope that the function of WT1 at the molecular level would become clearer. Although WT1 is known to be essential for kidney and gonad development, the mechanism by which it influences cell growth and differentiation is not well understood. There is considerable evidence suggesting that WT1 is a transcriptional regulator but this may not be its only role as it may have additional post-transcriptional functions. The complex nature of the gene, encoding 16 different isoforms, makes it particularly difficult to dissect WT1 function. Therefore if WT1-protein partners of known functions are found, then the molecular pathways in which WT1 is involved may start to fall into place. If protein partners are found whose function are not yet known, then studies could be done to try and identify their role before attempting to analyse the consequences of their interaction with WT1. Hopefully these studies may further our understanding of urogenital development and also how defects in the pathway might contribute to tumourigenesis.

7.1 A novel protein is identified which binds WT1 in the yeast two hybrid system

The yeast two hybrid system was used as a tool to identify protein-protein interactions. As discussed in chapter 3 WT1 $+/+$ was used to screen both an in-house M15 cDNA library, which had been made for the purpose of this study, and a Clontech foetal kidney library, in the hope of identifying proteins which bound to WT1. The screen with the M15 library only pulled out a few clones, some of which matched ESTs in the database. Unfortunately further analysis revealed that these positives were probably false. As the M15 library was not of a high quality (the insert sizes were small and the proportion of plasmids which were empty vectors was fairly high) it was decided to screen a commercial kidney library. This library was screened until it approached saturation. Clones were isolated which bound to WT1 but none of them originated from known proteins. No previously identified WT1 proteins were found. This could be because the particular library used does not contain cDNA clones representing those proteins or perhaps the library does not contain the full length clones, leading to lack of isolation if the clone does not contain the WT1-binding domain.

The Clontech screen isolated 5 clones that bound to WT1. Only one of these clones, clone D, bound to WT1 specifically, whereas the other 4 clones also bound to the negative control SNF1. Although these 4 clones could represent real WT1 interacting clones, they could equally be binding to WT1 due to a 'sticky' motif which binds to several proteins. No further investigations were carried out on these clones, which comprised out of frame fibrillin; 2 ESTs; and another clone which had no matches in the database. Clone D was chosen for further study as it bound to WT1 specifically from a range of factors tested and also bound to WT1 whether used as the bait or the prey, suggesting that the interaction detected between the two proteins is likely to be real. Clone D seemed to have a stronger association with the -KTS isoforms of WT1, and as domain binding studies showed, it interacted with the zinc finger region of WT1. As clone D was a novel WT1 binding protein, from then on it was called NAL (Novel Associating Ligand).

The NAL clone isolated was sequenced fully and intriguingly NAL matched the EST D14661 but only across the first 445 bases, after which there was a breakpoint where there was virtually no homology for the remainder of the sequence. This finding suggested that perhaps either NAL or the EST could have arisen from a chimaeric clone. Thus RT-PCR experiments were carried out, using primers that spanned the breakpoint, to determine which one of them was the true clone. It was shown that NAL could be amplified across the breakpoint in several cell lines used, whereas the EST could not, suggesting that the EST in the database had originated from a chimaeric clone. However closer examination of the breakpoint sequence of the EST suggested that there may be an exon-intron boundary and thus perhaps the EST in the database may have arisen from an unspliced pre-mRNA. This could have been verified by carrying out a PCR on genomic DNA. However as the EST also has a stop codon and a poly-adenylation site at its 3' end, perhaps the EST did not arise from a chimaeric clone or from unspliced pre-mRNA but could be an alternative transcript of NAL. If this hypothesis is correct, then it is surprising that the RT-PCR experiment carried out could not amplify the EST. However alternative splicing may be cell type dependent and so this transcript may not have been present in the particular cell lines used. Alternatively the transcript may be short lived or unstable in these cells. Although not tested, this hypothesis could have been checked by carrying out a Northern blot to see how many transcripts of NAL are present and their tissue distribution. Additionally RT-PCR could have been repeated using the original cell line from which the EST had been isolated. The C-terminus of NAL also matched ESTs in the database, but unfortunately a contig of the ESTs found did not extend sufficiently 5' to reach the breakpoint sequence.

The NAL protein sequence showed no homology to any proteins in the Swissprot database and the only motifs found within NAL were those of potential phosphorylation sites. NAL is glutamine rich and as glutamine rich domains have been found associated with transcription factors (Ladomery, 1997) this opens one avenue for future investigation (see section 7.6). There are no NAL homologues currently found in the invertebrate or fungi databases, suggesting that NAL, like WT1, is present only in vertebrates. The GenBank database contained a mouse EST homologous to NAL at its N-terminus. However, at the breakpoint, the mouse EST matched the human EST D14661 and not NAL, and thus a mouse homologue of NAL had to be isolated. A mouse library screen was performed to try to obtain the homologue of NAL. A Clontech 11 day mouse cDNA library was screened with the NAL probe and a clone was isolated which, when sequenced, was shown to be 94% identical at the nucleotide level to human NAL. Such high sequence conservation at the nucleotide level is not often seen, mainly because even if the amino acid sequence conservation is high, at the nucleotide level there can be wobble of the third base which frequently decreases the percentage identity to about 80%. This high nucleotide conservation in NAL suggests that there may be constraints at the RNA level, requiring the bases to be conserved. The fact that this mouse clone matches NAL, even after the breakpoint, supports the results which show that NAL was not produced from a chimaeric clone.

The human NAL clone isolated is not quite full length, but by comparing its sequence to the D14661 EST in the database, it is known that the start site is 9 base pairs further upstream. The isolated NAL clone is also only 14 base pairs short of reaching the termination site as compared to the EST found to match the 3' end of NAL. If a full length NAL clone were needed for future experiments, then the rapid amplification of cDNA ends (RACE) technique could be used to isolate an appropriate clone. The mouse Nal clone isolated is not full length either as it does not have a start site but it does contain the stop codon and the 3' UTR.

7.2 NAL binds to WT1 both *in vitro* and *in vivo*

The yeast two hybrid system identified NAL as potentially binding to WT1. This interaction had to be confirmed outside of the yeast two hybrid system to verify that the two proteins really bind to each other. Chapter 5 shows the glutathione-S-transferase (GST) *in vitro* assays used to provide additional evidence that WT1 and NAL can interact. It was found that even under high salt conditions, which disrupts complexes formed by weak interactions between two proteins, all four isoforms of *in*

vitro translated WT1 could bind to the GST-NAL fusion protein. Thus both the yeast two hybrid system and *in vitro* reactions show that WT1 and NAL can interact.

Despite these binding assays showing that the two proteins could bind, it was not known whether the interaction was physiologically relevant. In order to test this, *in vivo* co-immunoprecipitation assays were carried out. Transiently transfected HA-tagged NAL could not be co-immunoprecipitated with endogenous WT1, possibly because the levels of bound HA-NAL were too low to be detected. It was therefore necessary to obtain anti-NAL antibodies so that the experiments could be carried out using the endogenous protein.

A polyclonal antibody was raised against NAL by injecting GST-NAL fusion protein into rabbits. The antibody was shown to specifically recognise NAL on Western blots and so it could be used for the *in vivo* experiments. When the co-immunoprecipitations were carried out, this time using the endogenous proteins and detecting bound NAL with the polyclonal antibody, an interaction between NAL and WT1 was seen. As the relative intensities of the detected NAL was similar when the immunoprecipitation was carried out with anti-NAL antibody or with the anti-WT1 antibody it suggests that a high proportion of NAL is complexed with WT1. This experiment therefore conclusively demonstrated that in addition to binding in the yeast two hybrid system and *in vitro*, NAL and WT1 also bind *in vivo*.

7.3 Mapping the mouse and human NAL gene

The isolated mouse *Nal* cDNA was used to map the gene by FISH to mouse chromosome 17, near the centromere. This position was confirmed by using the EUCIB resource, which mapped *Nal* centromeric to the Brachyury (T) gene on chromosome 17. Interestingly, this region is very close to the T-locus and mutations within genes of the T-locus have profound effects on embryonal development (Bennett, 1975). So far no t-alleles have been found to map as close to the centromere of chromosome 17 as NAL. It is not therefore clear whether *Nal* is part of the T-locus and whether, by extrapolation, it plays an essential role in development.

The human NAL gene was mapped by FISH to chromosome 6q26-27. This region of human chromosome 6 shows conserved synteny with the centromeric region of murine chromosome 17, as shown by linkage maps. Interestingly this region of chromosome 6 has been associated with several malignancies and it is thought that perhaps a tumour suppressor gene may lie within this region (Theile et al., 1996).

7.4 NAL is ubiquitously expressed and may have alternative transcripts

The expression pattern of NAL was examined to see if, like WT1, it was tissue specific - being expressed mainly in the urogenital system - or whether it was more ubiquitously expressed. RT-PCR experiments carried out on adult mouse tissues showed that NAL was ubiquitously expressed in all tissues tested. Additionally the RT-PCR experiment amplified other bands smaller than the expected product in some tissues. Although these bands could be non-specific, this is unlikely as you might expect to see these non-specific bands in all samples. The nature of the shorter amplified bands was not investigated but it is possible that they could represent alternative transcripts of NAL. Future experiments should be carried out to determine whether NAL does have alternative transcripts. The RT-PCR experiments should be repeated and the products probed with an internal oligonucleotide of NAL to determine whether the shorter products are alternative transcripts of NAL, or alternatively the RT-PCR products could be directly cloned and sequenced.

Whole mount *in situ* hybridisation experiments were used to determine the spatial and temporal expression pattern of NAL. As presented in chapter 6, whole mount *in situ* hybridisations showed that NAL is ubiquitously expressed perhaps with increased expression seen in the tail extremity and the distal end of the limbs. The expression of NAL is already on at E=9.5 days (the youngest embryos examined), and remains on throughout the embryo at least until E=12.5d (the oldest embryos examined) – and during adulthood as the RT-PCR experiments showed. The expression pattern of NAL is therefore very different to that seen for WT1, as WT1 expression is localised to the urogenital ridge and the mesothelium and peaks at E=11.5d (Armstrong et al., 1993). By looking at the whole embryo, it was hard to tell whether the expression of NAL is increased in those areas that also express WT1, and so *in situ* hybridisation was carried out on histological sections where the internal regions of the embryo can be seen better. The preliminary data indicate that NAL is not expressed at higher levels in those areas that also express WT1, the staining pattern appearing uniform. These experiments also showed that, contrary to the results from the whole mounts, expression of NAL is not increased in the tail end or the distal end of the limbs, suggesting that these areas were more strongly stained in the whole mounts due to better accessibility of the probe.

The RT-PCR experiment and whole mount *in situ* hybridisations respectively showed that NAL was expressed in the adult kidney and in the developing urogenital ridge. It was of interest to determine whether during kidney development NAL, like WT1, was constrained to the developing nephros. Therefore immunohistochemistry

was carried out on E=17.5d kidneys which display all stages of nephrogenesis. The results showed that, unlike WT1 which is expressed only in the developing nephros, NAL is expressed throughout the kidney. The same was true in the testis, where NAL is ubiquitously expressed, while WT1 is confined to the Sertoli cells.

These expression studies therefore show that the pattern of NAL expression differs to that of WT1, and suggests that although NAL can associate with WT1, it can only do so in the tissues and cells where they are co-expressed. As NAL is ubiquitously expressed, it is possible that it may have a housekeeping role and that perhaps WT1 binds to NAL to somehow alter its function.

7.5 Cellular localisation and possible functions of NAL

Immunocytochemical analysis showed that NAL, like WT1, is totally nuclear. This is an important finding as if the two proteins are to interact then at least a portion of the proteins should be in the same compartment. Interestingly the nuclear staining pattern of NAL was shown to be speckled within a diffuse background. This is very similar to the nuclear distribution of WT1 (figure 6.5 and Larsson *et al.* 1995). The speckled structures of NAL were shown to overlap with the interchromatin granules which are reported to be sites where splicing factors are stored. This finding therefore suggested that NAL could have a role in splicing and so further tests were carried out to determine whether NAL behaved as a classical splicing factor.

Actinomycin D treatment was used to investigate whether, like classical splicing factors, NAL would re-distribute into large foci or alternatively whether NAL might behave like WT1 and U2AF65 and re-distribute into small foci localised around the remnants of the nucleolus (Larsson *et al.* 1995). The staining pattern of NAL did alter after actinomycin D treatment, becoming incorporated into several large foci as well as retaining some diffuse staining. These large foci overlapped with the large splice factor containing foci, and not with the smaller foci in which WT1 was found. Therefore NAL seems to be behaving as a more classical splicing factor in terms of nuclear distribution. However although it is tempting to speculate that NAL could have a role in splicing, caution should be exerted in the interpretation of these results because the speckles may contain proteins other than splicing factors. Nevertheless, if NAL and WT1 do play a role in splicing, and since they have been shown to interact *in vivo*, it is perhaps surprising that WT1 and NAL did not co-localise after actinomycin D treatment. However, since splicing does not occur after actinomycin D treatment, the results suggest that the interaction between WT1 and NAL could be important at the time when splicing occurs, but not during storage of the two proteins. Not much is known about the assembly of the spliceosome prior to

splicing. For instance it is not certain whether the spliceosome assembles in the speckles before being recruited to the site where it is required, or whether splice factors are recruited individually from the speckles as they are needed. Therefore if WT1 and NAL do have a role in splicing, they do not necessarily have to be interacting within the speckles, but only at the site where active splicing occurs.

Splicing requires several snRNPs and accessory proteins, whose interaction leads to the removal of introns by sequentially binding to the pre-mRNA. The events leading to fully spliced RNA include the stepwise assembly of the spliceosome via intermediate complexes as follows: in the E complex the U1 snRNP initially binds to the 5' splice site of the pre-mRNA and U2AF65 binds to the pyrimidine tract at the 3' splice site, both events requiring members of the SR family of splicing factors. The accessory factor U2AF65 then promotes binding of the U2 snRNP to the pre-mRNAs 3' branch point (Zamore et al., 1992) forming the A complex. The interaction between U1 and U2 brings the two splice sites together, committing the pre-mRNA to splicing. The trimeric complex of U5 and U4/U6 then completes the assembly of the spliceosome and this is known as the B complex. The splicing reaction is triggered by the release of U4 from the spliceosome complex, which probably acts to sequester U6 until it is required. The release of U4 now allows the RNA component of U6 to hybridise to part of the RNA within U2, the other part of U2 being hybridised to the branch point of the pre-mRNA. Following RNA rearrangement a transesterification occurs, cleaving the 5' exon from the intron producing a lariat structure. U5 may then have a role re-aligning the exons for the second transesterification event generating the mRNA.

The role of WT1 in splicing was examined by Davies *et al* (1998). They showed that WT1 could interact with U2AF65 *in vivo* and by doing so it may influence 3' splice site selection. They postulate that since WT1 has a putative RNA recognition motif (Kennedy et al., 1996), it may bind to the 3' splice site of the pre-RNA and present it to U2AF65, thus determining which 3' splice site is used. Since NAL localises to the speckles and because it interacts with WT1, NAL may also play a role in the 3' splice site selection. By binding to WT1, NAL perhaps stabilises the complex between WT1, U2AF65 and U2. This could be tested using the yeast two hybrid system, to see if NAL also binds to U2AF65. However, it has also been suggested that WT1 may act at several different stages during splicing (Davies et al., 1998) and so NAL could be interacting with WT1 at any of these points. There are several as yet unidentified non-snRNP splicing factors within the spliceosome and therefore NAL could be one of these, acting at any of the stages during splicing. Because NAL is ubiquitously expressed, it is plausible that it is part of the spliceosome and

that perhaps WT1 binds to NAL to somehow affect the outcome of the overall splicing event in a tissue specific manner.

7.6 Future directions

Although this study has not managed to find a WT1-binding protein which directly helps to further our understanding of WT1, it has identified a novel protein, NAL, which binds to WT1 *in vivo*. Several speculative functions of NAL have been proposed but its true role and the consequences of its interaction with WT1 remains to be investigated.

Does NAL play a role in splicing?

The studies carried out in chapter 6 showed the nuclear distribution of NAL. It was shown that NAL had a speckled appearance within a diffuse background and that the speckles overlapped with the speckles associated with splicing factors. This nuclear distribution of NAL is thus suggestive of a role in splicing, which could be studied in more detail.

One way of determining whether NAL acts like a splicing factor is by looking at its dynamic properties. The movement of splicing factors within the nucleus has been looked at and seen to leave the speckles in peripheral extensions and accumulate at sites of active transcription (Misteli et al.1997). This movement of splicing factors can be blocked when transcription inhibitors are used. Thus to investigate whether NAL can move and behave in a splice factor-like manner, a GFP-NAL fusion protein can be made and the dynamics of NAL visualised in living cells.

Immunofluorescent studies under a variety of different conditions have provided preliminary ways of classifying splice factors. For example RNase A treatment of cells affects snRNPs and non-snRNP splice factors differently. After treatment with RNase A, snRNPs leave the speckles and adopt a diffuse localisation whereas non-snRNPs stay in the speckles which become enlarged. This difference in behaviour is probably due to snRNP particles requiring RNA-RNA or RNA-protein interactions to be associated with the speckled region whereas non-snRNP splice factors require protein-protein interactions (Spector et al., 1991). Thus these experiments could also be carried out for NAL and they may show what interactions are keeping NAL in the speckles and whether NAL is acting more like an snRNP or non-snRNP splice factor. Alternatively cells could be micro-injected with antisense oligonucleotides. The assembly of snRNPs on pre-mRNA is dependent on several RNA-RNA interactions. Thus antisense oligonucleotides which disrupts such interactions, like those which are complementary to the 5' end of the U1 snRNA, can be injected into

cells disrupting the spliceosomes (Larsson et al., 1995). Following this treatment the snRNPs assemble in large foci with little diffuse staining. If the sub-nuclear localisation of NAL is dependent on spliceosome assembly, then NAL should also redistribute into these large foci. Although this type of immunofluorescence analysis can give important clues as to potential function, ultimately more direct assays would be required to determine whether NAL is part of the spliceosome.

To ascertain whether NAL is part of the spliceosome, co-immunoprecipitations could be carried out. If NAL is part of the spliceosome, then the 474 anti-NAL antibody should be able to immunoprecipitate spliceosomes assembled *in vitro*. If NAL is found to be associated with spliceosomes, then immunodepletion experiments could also be carried out, where nuclear extracts would be incubated with the NAL antibody to remove NAL. These NAL-depleted extracts could then be tested for their ability to splice exogenously added pre-mRNA. If splicing cannot occur, it would be suggestive of an essential role for NAL in splicing. Purified NAL could be added back to the NAL-depleted extracts and if NAL is in some way involved in splicing, the splicing activity may be restored.

If the previous assays all point towards NAL being a splicing factor, then the role of NAL in splicing could be investigated. The postulated involvement of WT1 at the 3' splice site makes an analysis of NAL at similar sites particularly appealing. To determine whether NAL is present in early stages of spliceosome assembly, biotinylated pre-mRNA containing only a 5' or a 3' splice site could be prepared. Following incubation with nuclear extracts from NAL expressing cells, streptavidin beads are used to purify the pre-mRNA. A Western blot is then carried out to investigate whether NAL is incorporated into the earliest splicing complexes present exclusively at the 5' or 3' splice sites. To study whether NAL affects splice site selection during alternative splicing, splicing assays using reporter genes can be carried out. Minigenes containing different alternative splice sites can be transfected into cells either with or without NAL. The harvested RNA is then used for RT-PCR and according to the products of the PCR, the effect of NAL can be determined. Similar assays could be carried out to see if NAL has any antagonistic effects on splicing factors known to affect splice site selection. The yeast two hybrid method and immunoprecipitation assays could be carried out to determine with which components of the splicing machinery NAL is interacting, and thus start gaining some idea as to which splicing events NAL is involved in.

Does NAL play a role in transcription?

The immunocytochemistry showed that apart from being present in the speckles, NAL also had a diffuse staining pattern. This similar staining pattern is seen for WT1, and probably represents those isoforms which lack the KTS insert and whose main role is proposed to be transcriptional (Larsson et al., 1995). As NAL has a glutamine rich, a motif commonly found in transcription regulators, then perhaps a potential role of NAL in transcription should be investigated.

Transcription factors typically have two domains, a DNA binding domain and an activation domain. The motif search within NAL did not reveal any DNA-binding domains, but this does not exclude the possibility that NAL could perhaps bind to DNA especially if it does play a role in transcription. Thus a general DNA binding assay could be carried out using DNA-cellulose columns to see whether NAL is able to bind to DNA. If NAL is found to bind to DNA, then the specific sequence it binds to could be investigated by using DNA selection assays. Band shift assays could also be carried out to verify that NAL interacts directly with the isolated target.

If NAL does not bind to DNA directly then it could still be acting as a co-activator or repressor. The role of NAL in transcriptional control could be examined by fusing NAL to the yeast Gal4 DNA binding domain. The construct would then be co-transfected into cells with a Gal4-dependent reporter, and activation or repression of the reporter gene looked at. It may be that NAL itself is not an activator or repressor of transcription, but perhaps it could affect the transcriptional activity of WT1. The yeast two hybrid system showed that NAL bound better to the -KTS isoform of WT1 and so it may do so to affect the transcriptional activity of WT1. Transcriptional reporter assays could therefore be used to determine whether NAL has any effect on the transcriptional activity of WT1. If NAL itself has a transcriptional activity then the effect of WT1 on NAL could also be looked at.

Does NAL play a role in proliferation?

As mentioned previously, the whole mount *in situ* hybridisation data showed a tenuous link between NAL and proliferation, hence experiments could be carried out to determine whether NAL may play a role in cell cycle regulation. The levels of NAL could be looked at during the cycle to see if they are altered at any particular stage. Of course it could be that the levels of NAL itself are not altered but its activity is. Alternatively NAL could be overexpressed in cells to see if this causes any effect – either promoting growth or apoptosis. The importance of the WT1 function in suppression of cell growth was shown when WT1 was identified as a tumour suppressor gene. In these experiments, wild type *WT1* was reintroduced into

a Wilms' tumour cell line expressing an aberrantly spliced *WT1* transcript, and shown that it could suppress cell growth (Haber et al., 1993). It would be of interest to determine whether NAL itself stimulates or inhibits growth, or whether it interferes with this aspect of WT1 function.

Does NAL play a role in development or disease?

The studies carried out in chapter 4 mapped the mouse *Nal* gene to chromosome 17, centromeric to the Brachyury gene. *Nal* therefore maps very close to the T-locus. Mutations within genes of the T-locus have profound effects on embryonal development and as it is not yet clear whether *Nal* is part of the T-locus and therefore whether it plays an essential role in development, it would be of interest to make a *Nal* null mouse as the phenotype of such a mouse may help in understanding the role of NAL.

The chromosomal mapping of human NAL is suggestive of a role in disease. NAL was mapped to chromosome 6q26-27, a region which has been associated with several malignancies and which is thought to contain a tumour suppressor gene. Future work could therefore include looking at renal cell carcinomas and ovarian cancers which have been shown to have deletions in this region of chromosome 6 to see if NAL is perhaps deleted or mutated or whether its expression is altered compared to normal tissue. These experiments might help to determine whether NAL plays a role in these malignancies. It would also be of interest to see if the expression of NAL is altered in Wilms' tumours compared to normal tissue, as NAL may be regulated by WT1.

Does NAL have alternative transcripts?

Future work should also concentrate on determining whether NAL does have alternative transcripts as indicated by both RT-PCR experiments and the EST database matches. Initially the RT-PCR experiments ought to be repeated and the products produced should be cloned and sequenced. If these products do seem to originate from alternative transcripts, then perhaps a cDNA library should be screened with NAL probes to see whether alternative transcripts can be isolated. A genomic library could also be screened to be able to analyse the genomic organisation of NAL and by doing so being able to follow which exons the alternative transcripts contain. If alternative transcripts are found then it would be of interest to know whether they carry out different roles, just like the different WT1 isoforms are proposed to do.

Is NAL part of a network?

Apart from the aforementioned ways of studying the role of NAL, perhaps the yeast two hybrid screen could be used to identify NAL-binding proteins. In this way a network of interactions between NAL, WT1 and other cellular proteins can start to be built up so as to shed some light on the cellular pathways in which NAL and WT1 are involved.

This future work should therefore help in trying to understand the role of this novel protein and the reason behind its interaction with WT1.

Bibliography

Bibliography

- Adachi, Y., Matsubara, S., Pedraza, C., Ozawa, M., Tsutsui, J.-I., Takamatsu, H., Noguchi, H., Akiyama, T., and Muramatsu, T. (1996). Midkine as a novel target gene for the Wilms' Tumor suppressor gene (WT1). *Oncogene* 13, 2197-2203.
- Allen, J. B., Walberg, M. W., Edwards, M. C., and Elledge, S. J. (1995). Finding prospective partners in the library: the two-hybrid system and phage display find a match. *Trends in Biochemical Sciences* 20, 511-516.
- Armstrong, J. F., Pritchard Jones, K., Bickmore, W. A., Hastie, N. D., and Bard, J. B. (1993). The expression of the Wilms' tumour gene, WT1, in the developing mammalian embryo. *Mechanisms of Development* 40, 85-97.
- Barboux, S., Niaudet, P., Gubler, M.-C., Grunfeld, J.-P., Jaubert, F., Kuttann, F., Fekete, C. N., Souleyreau-Therville, N., Thibaud, E., Fellous, M., and McElreavey, K. (1997). Donor splice-site mutations in WT1 are responsible for Frasier syndrome. *Nature Genetics* 17, 467-470.
- Bardeesy, N., and Pelletier, J. (1998). Overlapping RNA and DNA binding domains of the wt1 tumor suppressor gene product. *Nucleic Acids Research* 26, 1784-1792.
- Bartolomei, M., Zemel, S., and Tilghman, S. M. (1991). Parental imprinting of the mouse H19 gene. *Nature* 351, 153-155.
- Beckwith, J. B. (1969). Macroglossa, omphalocoele, adrenal cytomegaly, gigantism and hyperplastic visceromegaly. *Birth Defects* 5, 188-196.
- Beckwith, J. B. (1997). New Developments in the Pathology of Wilms' Tumor. *Cancer Investigation* 15, 153-162.
- Beckwith, J. B., Kiviat, N. B., and Bonadio, J. F. (1990). Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. *Pediatric Pathology* 10, 1-36.
- Bender, L., Shuen Lo, H., Kokojan, V., and Peterson, J. (1996). Associations among PH and SH3 domain-containing proteins and Rho-type GTPases in yeast. *The Journal of Cell Biology* 133, 879-894.
- Bennett, D. (1975). The T-locus of the mouse. *Cell* 6, 441-454.
- Bickmore, W. A., Oghene, K., Little, M. H., Seawright, A., van Heyningen, V., and Hastie, N. D. (1992). Modulation of DNA binding specificity by alternative splicing of the Wilms' tumour WT1 gene transcript. *Science* 257, 235-237.
- Brenner, B., Wildhardt, G., Schneider, S., and Royer-Pokora, B. (1992). RNA polymerase chain reaction detects different levels of four alternatively spliced WT1 transcripts in Wilms' tumour. *Oncogene* 7, 1431-1433.
- Breslow, N., and Beckwith, J. B. (1988). Epidemiological features of Wilms' tumour: Results of the National Wilms' tumour study. *J.Nat.Cancer Inst* 68, 429-436.
- Breslow, N., Olson, J., Moksness, J., Beckwith, J. B., and Grundy, P. (1996). Familial Wilms' tumor: A descriptive study. *Medical and Pediatric Oncology* 27, 398-403.
- Bruening, W., Gros, P., Sato, T., Stanimir, J., Nakamura, Y., Housman, D., and Pelletier, J. (1993). Analysis of the 11p13 Wilms' tumor suppressor gene (WT1) in ovarian tumors. *Cancer Investigation* 11, 393-399.
- Bruening, W., and Pelletier, J. (1996). A non-AUG translational initiation event generates novel WT1 isoforms. *Journal of Biological Chemistry* 271, 8646-8654.

Buckler, A. J., Pelletier, J., Haber, D. A., Glaser, T., and Housman, D. E. (1991). Isolation, characterization, and expression of the murine Wilms' tumor gene (WT1) during kidney development. *Molecular and Cellular Biology* 11, 1707-1712.

Caceres, J. F., Misteli, T., Sreaton, G. R., Spector, D. L., and Krainer, A. R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *Journal of Cell Biology* 138, 225-238.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C., and Housman, D. E. (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* 60, 509-520.

Caricasole, A., Duarte, A., Larsson, S. H., Hastie, N. D., Little, M., Holmes, G., Todorov, I., and Ward, A. (1996). RNA binding by the Wilms Tumour suppressor zinc finger proteins. *Proc.Natl.Acad.Sci.U.S.A.* 93, 7562-7566.

Carmo-Fonseca, M., Pepperkok, R., Carvalho, M. T., and Lamond, A. I. (1992). Transcription-Dependent Colocalization of the U1, U2, U4/U6, and U5 Snrnps in Coiled Bodies. *Journal of Cell Biology* 117, 1-14.

Celenza, J. L., Eng, F. J., and Carlson, M. (1989). Molecular analysis of the SNF4 gene of *Saccharomyces cerevisiae*: Evidence for physical association of the SNF4 protein with the SNF1 protein kinase. *Molecular and Cellular Biology* 9, 5045-5054.

Charlier, J.-P., Larsson, S., Miyagawa, K., van Heyningen, V., and Hastie, N. D. (1995). Does the Wilms' Tumour suppressor gene, WT1, play roles in both splicing and transcription? *The Journal of Cell Science* 19 (supplement), 95-99.

Cohen, H. T., Bossone, S. A., Zhu, G., McDonald G.A., and Sukhatme, V. P. (1997). Sp1 is a critical regulator of the *Wilms' tumor-1* gene. *The Journal of Biological Chemistry* 272, 2901-2913.

Compton, D. A., Weil, M. M., Jones, C., Riccardi, V. M., Strong, L. C., and Saunders, G. F. (1988). Long range physical map of the Wilms' tumor-aniridia region on human chromosome 11. *Cell* 55, 827-836.

Cooke, I. E., Shelling, A. N., Le Meuth, V. G., Charnock, M. L., and Ganesan, T. S. (1996). Allele loss on chromosome arm 6q and fine mapping of the region at 6q27 in epithelial ovarian cancer. *Genes Chromosomes Cancer* 15, 223-233.

D'Angio, G. J., Breslow, N., Beckwith, B., Evans, A., Baum, E., DeLorimier, A., Fernbach, D., Habovsky, E., Jones, B., Kelalis, P., Othersen, B., Tefft, M., and Thomas, P. R. M. (1989). Treatment of Wilm's Tumour. *Cancer* 64, 349-360.

Davies, J. A. (1996). Mesenchyme to epithelium transition during development of the mammalian kidney tubule. *Acta Anatomica* 156, 187-201.

Davies, J. A., and Bard, J. B. L. (1996). Inductive interactions between the mesenchyme and the ureteric bud. *Experimental Nephrology* 4, 77-85.

Davies, J. A., and Brandli, A. W. (1994). The kidney development database: <http://www.ana.ed.ac.uk/anatomy/database/kidbase/kidhome/html>.

Davies, R. C., Calvio, C., Bratt, E., Larsson, S. H., Lamond, A. I., and Hastie, N. D. (1998). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. *Genes & Development* 12, 3217-3225.

Davis, L. M., Stallard, R., Thomas, G. H., Coullin, P., Junien, C., Nowak, N. J., and Shows, T. B. (1988). Two anonymous DNA segments distinguish the Wilms' tumour and aniridia locus. *Science* 241, 840-842.

DeChiara, T. M., Robertson, E. J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 64, 849-859.

- Dehbi, M., Ghahremani, M., Lechner, M., Dressler, G., and Pelletier, J. (1996). The paired-box transcription factor, Pax2, positively modulates expression of the Wilms' tumor suppressor gene (WT1). *Oncogene* 13, 447-453.
- Dehbi, M., and Pelletier, J. (1996). Pax8 mediated activation of the wt1 tumour suppressor gene. *EMBO journal* 15, 4297-4306.
- Denys, P., Malvaux, P., van de Berghe, H., Tanghe, W., and Proesmans, W. (1967). Association d'un syndrome anatomo-pathologique de pseudohermaphrodisme masculin, d'une tumeur de Wilms, d'une néphropathie parenchymateuse et d'une mosaïcisme XX/XY. *Arch.Franc.Ped.* 24, 729-739.
- Dey, B. R., Sukhatme, V. P., Roberts, A. B., Sporn, M. B., Rauscher, F. J., and Kim, S. J. (1994). Repression of the Transforming Growth-Factor-Beta-1 gene by the Wilms-Tumor suppressor Wt1 gene-product. *Molecular Endocrinology* 8, 595-602.
- Drash, A., Sherman, F., Hartman, W. H., and Blizzard, R. M. (1970). A syndrome of psuedohermaphrodism, Wilms' tumour, hypertension, and degenerative renal disease. *Journal of Pediatrics* 76, 585-593.
- Drummond, I. A., Madden, S. L., Rohwer Nutter, P., Bell, G. I., Sukhatme, V. P., and Rauscher, F. J. (1992). Repression of the insulin-like growth factor II gene by the Wilms tumor suppressor WT1. *Science* 257, 674-678.
- Durfee, T., Becherer, K., Chen, P., Yeh, S., Yang, Y., Kilburn, A., Lee, W., and Elledge, S. J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes and Development* 7, 555-569.
- Edwards-Gilbert, G., Veraldi, K. L., and Milcarek, C. (1997). Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Research* 25, 2547-2561.
- Englert, C., Hou, X., Maheswaran, S., Bennett, P., Ngwu, C., Re, G. G., Garvin, A. J., Rosner, M. R., and Haber, D. A. (1995a). WT1 suppresses synthesis of the epidermal growth factor receptor and induces apoptosis. *EMBO Journal* 14, 4662-4675.
- Englert, C., Maheswaran, S., Garvin, A. J., Kreidberg, J., and Haber, D. A. (1997). Induction of p21 by the Wilms' tumor suppressor gene WT1. *Cancer Research* 57, 1429-1434.
- Englert, C., Vidal, M., Maheswaran, S., Ge, M., Ezzell, R. M., Isselbacher, K. J., and Haber, D. A. (1995b). Truncated WT1 mutants alter the subnuclear localization of the wild-type protein. *Proc.Natl.Acad.Sci.U.S.A.* 92, 11960-11964.
- Estojak, J., Brent, R., and Golemis, E. A. (1995). Correlation of two-hybrid affinity data with in vitro measurements. *Molecular and Cellular Biology* 15, 5820-5829.
- Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* 340, 245-246.
- Fraizer, G. C. (1995). Expression of the tumour-suppressor gene WT1 in both human and mouse bone marrow. *Blood* 86, 4704-4706.
- Fraizer, G. C., Wu, Y. L., Hewitt, S. M., Maity, T., Ton, C. T. T., Huff, V., and Saunders, G. F. (1994). Transcriptional regulation of the human Wilms' tumor gene (WT1). *The Journal of Biological Chemistry* 269, 8892-8900.
- Francke, U., Holmes, L. B., Atkins, L., and Riccardi, V. M. (1979). Aniridia-Wilms' tumour associations: Evidence for specific deletion of 11p13. *Cytogenet.Cell Genet.* 24, 185-192.
- Fromont-Racine, M., Rain, J.-C., and Legrain, P. (1997). Towards a functional analysis of the yeast genome through exhaustive two hybrid screens. *Nature Genetics* 16, 277-282.

- Gama-Carvalho, M., Krauss, R. D., Chiang, L. J., Valcarcel, J., Green, M. R., and Carmo-Fonseca, M. (1997). Targeting of U2AF65 to sites of active splicing in the nucleus. *Journal of Cell Biology* 137, 975-987.
- Gessler, M., Konig, A., Arden, K., Grundy, P., Orkin, S., Sallan, S., Peters, C., Ruyle, S., Mandell, J., Li, F., Cavenee, W., and Bruns, G. (1994). Infrequent mutation of the WT1 gene in 77 Wilms' tumors. *Human Mutation* 3, 212-222.
- Gessler, M., Konig, A., and Bruns, G. A. (1992). The genomic organization and expression of the WT1 gene. *Genomics* 12, 807-813.
- Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H., and Bruns, G. A. (1990). Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. *Nature* 343, 774-778.
- Goodyer, P., Dehbi, M., Torban, E., Bruening, W., and Pelletier, J. (1995). Repression of the retinoic acid receptor-alpha gene by the Wilms-Tumor suppressor gene-product, Wt1. *Oncogene* 10, 1125-1129.
- Grundy, P., Koufos, A., Morgan, K., Li, F. P., Meadows, A. T., and Cavenee, W. K. (1988). Familial predisposition to Wilms' tumour does not map to the short arm of chromosome 11. *Nature* 336, 374-376.
- Grundy, P. E., Telzerow, P. E., Breslow, N., Moskess, J., Huff, V., and Paterson, M. C. (1994). Loss of heterozygosity for chromosomes 16q and 1p in Wilms' tumors predicts an adverse outcome. *Cancer Research* 54, 2331-2333.
- Haber, D. A., Park, S., Maheswaran, S., Englert, C., Re, G. G., Hazen Martin, D. J., Sens, D. A., and Garvin, A. J. (1993). WT1-mediated growth suppression of Wilms tumor cells expressing a WT1 splicing variant. *Science* 262, 2057-2059.
- Haber, D. A., Sohn, R. L., Buckler, A. J., Pelletier, J., Call, K. M., and Housman, D. E. (1991). Alternative splicing and genomic structure of the Wilms tumor gene WT1. *Proc.Natl.Acad.Sci.U.S.A.* 88, 9618-9622.
- Hamilton, T. B., Barilla, K. C., and Romaniuk, P. J. (1995). High affinity binding sites for the Wilms' tumour suppressor protein WT1. *Nucleic Acids Research* 23, 277-284.
- Hammond, K. L., Hanson, I. M., Brown, A. G., Lettice, L. A., and Hill, R. E. (1998). Mammalian and *Drosophila dachshund* genes are related to the Ski proto-oncogene and are expressed in eye and limb. *Mechanisms of Development* 74, 121-131.
- Harrington, M. A., Konicek, B., Song, A., Xia, X. L., Fredericks, W. J., and Rauscher, F. J. (1993). Inhibition of colony-stimulating factor-1 promoter activity by the product of the Wilms' tumor locus. *The Journal of Biological Chemistry* 268, 21271-21275.
- Hastie, N. D. (1994). The genetics of Wilms' tumour - A case of disrupted development. *Annual Review of Genetics* 28, 523-558.
- Henry, I., Bonaiti-Pellie, C., Chehensse, V., Beldjord, C., Schwartz, C., Utermann, G., and Junien, C. (1991). Uniparental paternal disomy in a genetic cancer-predisposing syndrome. *Nature* 351, 665-667.
- Henry, I., Grandjouan, S., Couillin, P., Barichard, F., Huerre-Jeanpierre, C., Glaser, T., Philip, T., Lenoir, G., Chaussain, J. L., and Junien, C. (1989). Tumor-specific loss of 11p15.5 alleles in del 11p13 Wilms tumor and in familial adrenocortical carcinoma. *Proc.Natl.Acad.Sci.U.S.A.* 86, 3247-3251.
- Hewitt, S. M., Hamada, S., McDonnell, T. J., Rauscher III, F. J., and Saunders, G. F. (1995). Regulation of the proto-oncogenes bcl-2 and c-myc by the Wilms' tumor suppressor gene WT1. *Cancer Research* 55, 5386-5389.

- Hofmann, W., Royer, H. D., Drechsler, M., Schneider, S., and Royer-Pokora, B. (1993). Characterization of the transcriptional regulatory region of the human WT1 gene. *Oncogene* 8, 3123-3132.
- Hsu, S. Y., Kubo, M., Chun, S. Y., Haluska, F. G., Housman, D. E., and Hsueh, A. J. W. (1995). Wilms' tumor protein WT1 as an ovarian transcription factor: decreases in expression during follicle development and repression of inhibin alpha gene promoter. *Molecular Endocrinology* 9, 1356-1366.
- Hsu, T., Gogos, J. A., Kirsh, S. A., and Kafatos, F. C. (1992). Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene. *Science* 257, 1946-1950.
- Huff, V., Amos, C. I., Douglass, E. C., Fischer, R., Geiser, C. F., Krill, C. E., Strong, L. C., and McDonald, J. M. (1997). Evidence for genetic heterogeneity in familial Wilms' tumor. *Cancer Research* 57, 1859-1862.
- Huff, V., Campton, D. A., Chao, L. Y., Strong, L. C., Geiser, G. F., and Saunders, G. F. (1988). Lack of linkage of familial Wilms' tumour to chromosomal band 11p13. *Nature* 336, 377-378.
- Jackson, I. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acid Research* 19, 3795-3798.
- Johnstone, R. W., See, R. H., Sells, S. F., Wang, J., Muthukkumar, S., Englert, C., Haber, D. A., Licht, J. D., Sugrue, S. P., Roberts, T., Rangnekar, V. M., and Shi, Y. (1996). A novel repressor, par-4, modulates transcription and growth suppression functions of the Wilms' tumor suppressor WT1. *Molecular and Cellular Biology* 16, 6945-6956.
- Johnstone, R. W., Wang, J., Tommerup, N., Vissing, H., Roberts, T., and Shi, Y. (1998). Cio 1 is a novel WD40 protein that interacts with the tumour suppressor protein WT1. *The Journal of Biological Chemistry* 273, 10880-10887.
- Karnik, P., Chen, P., Paris, M., Yeger, H., and Williams, B. R. G. (1998). Loss of heterozygosity at chromosome 11p15 in Wilms tumors: identification of two independent regions. *Oncogene* 17, 237-240.
- Kennedy, D., Ramsdale, T., Mattick, J., and Little, M. (1996). An RNA recognition motif in Wilms' tumour protein (WT1) revealed by structural modelling. *Nature Genetics* 12, 329-332.
- Kent, J., Coriat, A. M., Sharpe, P. T., Hastie, N. D., and van Heyningen, V. (1995). The evolution of WT1 sequence and expression pattern in vertebrates. *Oncogene* 11, 1781-1792.
- King-Underwood, L., Renshaw, J., and Pritchard-Jones, K. (1996). Mutations in the Wilms' tumour gene WT1 in leukemias. *Blood* 87, 2171-2179.
- Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research* 12, 857-872.
- Knudson, A. G. (1971). Mutation and cancer: Statistical study of retinoblastoma. *Proc Natl Acad Sci USA* 68, 820-823.
- Knudson, A. G., and Strong, L. C. (1972). Mutation and cancer: a model for Wilms' tumour of the kidney. *Journal of National Cancer Investigation* 48, 313-324.
- Koseki, C., Herzlinger, D., and Al-Awqati, Q. (1992). Apoptosis in metanephric development. *Journal of Cell Biology* 119, 1327-1333.
- Koufos, A., Grundy, P., Morgan, K., Aleck, K. A., and Hadro, T. (1989). Familial Wiedemann-Beckwith syndrome and a second Wilms' tumour locus both map to 11p15.5. *American Journal of Human Genetics* 44, 711-719.
- Kramer, S., Meadows, A. T., and Jarrett, P. (1984). Racial variation in incidence of Wilms' tumor: relationship to congenital anomalies. *Medical and Paediatric Oncology* 12, 401-405.

- Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell* 74, 679-691.
- Kudoh, T., Ishidate, T., Moriyama, M., Toyoshima, K., and Akiyama, T. (1995). G1 phase arrest induced by Wilms tumor protein WT1 is abrogated by cyclin/CDK complexes. *Proc.Natl.Acad.Sci.U.S.A.* 92, 4517-4521.
- Ladomery, M. (1997). Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *Bioessays* 19, 903-909.
- Langerak, A. W., Williamson, K. A., Miyagawa, K., Hagemijer, A., Versnel, M. A., and Hastie, N. D. (1995). Expression of the Wilms' tumor gene WT1 in human malignant mesothelioma cell lines and relationship to platelet-derived growth factor A and insulin-like growth factor 2 expression. *Genes Chromosomes and Cancer* 12, 87-96.
- Larsson, S. H., Charlieu, J.-P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V., and Hastie, N. D. (1995). Subnuclear localization of WT1 in splicing or transcription domains is regulated by alternative splicing. *Cell* 81, 391-401.
- Lechner, M. S., and Dressler, G. R. (1997). The molecular basis of embryonic kidney development. *Mechanisms of Development* 62, 105-120.
- Lee, H., and Maible, N. J. (1998). Isolation and characterization of four alternate c-erbB3 transcripts expressed in ovarian carcinoma-derived cell lines and normal human tissues. *Oncogene* 16, 3243-3252.
- Lesage, P., Yang, X., and Carlson, M. (1994). Analysis of the SIP3 protein identified in a two hybrid screen for interaction with the SNF1 protein kinase. *Nucleic Acids Research* 22, 597-603.
- Lewis, W. H., Yeger, H., Bonetta, L., Chan, H. S., Kang, J., Junien, C., Cowell, J., Jones, C., and Dafoe, L. A. (1988). Homozygous deletion of a DNA marker from chromosome 11p13 in sporadic Wilms tumor. *Genomics* 3, 25-31.
- Li, H., and Bingham, P. M. (1991). Arginine/serine -rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* 67, 335-342.
- Little, M., Holmes, G., Bickmore, W., van Heyningen, V., Hastie, N., and Wainwright, B. (1995). DNA binding capacity of the WT1 protein is abolished by Denys-Drash syndrome WT1 point mutations. *Human Molecular Genetics* 4, 351-358.
- Little, M., and Wells, C. (1997). A clinical overview of *WT1* gene mutations. *Human Mutation* 9, 209-225.
- Lustig-Yariv, O., Schulze, E., Komitowski, D., Erdmann, V., Schneider, T., de Groot, N., and Hochberg, A. (1997). The expression of the imprinted genes H19 and IGF-2 in choriocarcinoma cell lines. Is H19 a tumor suppressor gene? *Oncogene* 15, 169-177.
- Madden, S. L., Cook, D. M., Morris, J. F., Gashler, A., Sukhatme, V. P., and Rauscher III, F. J. (1991). Transcriptional repression mediated by the WT1 Wilms tumor gene product. *Science* 253, 1550-1553.
- Madden, S. L., Cook, D. M., and Rauscher, F. J. (1993). A structure-function analysis of transcriptional repression mediated by the WT1, Wilms' tumor suppressor protein. *Oncogene* 8, 1713-1720.
- Maheswaran, S., Englert, C., Bennett, P., Heinrich, G., and Haber, D. A. (1995). The WT1 gene product stabilizes p53 and inhibits p53-mediated apoptosis. *Genes and Development* 9, 2143-2156.

- Maheswaran, S., Englert, C., Lee, S. B., Ezzle, R. M., Settleman, J., and Haber, D. A. (1998a). E1B 55K sequesters WT1 along with p53 within a cytoplasmic body in adenovirus-transformed kidney cells. *Oncogene* 16, 2041-2050.
- Maheswaran, S., Englert, C., Zheng, G., Lee, S. B., Wong, J., Harkin, D. P., Bean, J., Ezzell, R., Garvin, A. J., McCluskey, R. T., DeCaprio, J. A., and Haber, D. A. (1998b). Inhibition of cellular proliferation by the Wilms tumour suppressor WT1 requires association with the inducible chaperone Hsp70. *Genes & Development* 12, 1108-1120.
- Maheswaran, S., Park, S., Bernard, A., Morris, J. F., Rauscher III, F. J., Hill, D. E., and Haber, D. A. (1993). Physical and functional interaction between WT1 and p53 proteins. *Proc.Natl.Acad.Sci.U.S.A.* 90, 5100-5104.
- Martinerie, C., Chevalier, G., Rauscher, F. J., and Perbal, B. (1996). Regulation of nov by WT1: a potential role for nov in nephrogenesis. *Oncogene* 12, 1479-1492.
- Matsunaga, E. (1981). Genetics of Wilms' Tumor. *Human Genetics* 57, 231-246.
- Maw, M. A., Grundy, P. E., Millow, L. J., Eccles, M. R., and Dunn, R. S. (1992). A Third Wilms' tumour locus on chromosome 16q. *Cancer Research* 52, 3094-3098.
- McDonald, J. M., Douglass, E. C., Fisher, R., Geiser, C. F., Krill, C. E., Strong, L. C., Virshup, D., and Huff, V. (1998). Linkage of familial Wilms' Tumor predisposition to chromosome 19 and a two-locus model for the etiology of familial tumors. *Cancer Research* 58, 1387-1390.
- McLaren, A. (1998). Gonad development: assembling the mammalian testis. *Current Biology* 8.
- Menasce, L. P., Orphanos, V., Santibanez-Koref, M., Boyle, J. M., and Harrison, C. J. (1994). Common region of deletion on the long arm of chromosome 6 in non-Hodgkin's lymphoma and acute lymphoblastic leukaemia. *Genes Chromosomes and Cancer* 10, 286-288.
- Menke, A., McInnes, L., Hastie, N. D., and Schedl, A. (1998a). The Wilms' tumor suppressor WT1: Approaches to gene function. *Kidney International* 53, 1512-1518.
- Menke, A. L., Riteco, N., van Ham, R. C., De Bruyne, C., Rauscher III, F. J., Van Der Eb, A. J., and Jochemsen, A. G. (1996). Wilms' tumour 1 splice site variants have opposite effects on the tumorigenicity of adenovirus-transformed baby-rat kidney cells. *Oncogene* 12, 537-546.
- Menke, A. L., van der Eb, A. J., and Jochemsen, A. G. (1998b). The Wilms' tumor 1 gene: oncogene or tumor suppressor gene? *International Review of Cytology* 181, 151-212.
- Miles, C., Elgar, G., Coles, E., Kleinjan, D.-J., van Heyningen, V., and Hastie, N. (1998). Complete sequencing of the *Fugu* WAGR region from WT1 to PAX6 - dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc.Natl.Acad.Sci.U.S.A* 95, 13068-13072.
- Miller, R. W., Fraumeni, J. F., and Manning, M. D. (1964). Association of Wilms' tumour with aniridia, hemihypertrophy and other abnormalities. *New England Journal of Medicine* 270, 922-927.
- Misteli, T., Caceres, J. F., and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* 387, 523-527.
- Misteli, T., and Spector, D. L. (1998). The cellular organization of gene expression. *Current Opinion in Cell Biology* 10, 323-331.
- Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245, 371-378.
- Miwa, H., Beran, M., and Saunders, G. F. (1992). Expression of the Wilms' tumor gene (WT1) in human leukemias. *Leukemia* 6, 405-409.
- Miyagawa, K., Kent, J., Moore, A., Charlier, J.-P., Little, M. H., Williamson, K. A., Kelsey, A., Brown, K. W., Hassam, S., Briner, J., Hayashi, Y., Hirai, H., Yazaki, Y., van Heyningen, V., and

- Hastie, N. D. (1998). Loss of WT1 function leads to ectopic myogenesis in Wilms' tumour. *Nature Genetics* 18, 15-17.
- Moffett, P., Bruening, W., Nakagama, H., Bardeesy, N., Housman, D., Houseman, D. E., and Pelletier, J. (1995). Antagonism of WT1 activity by protein self-association. *Proc. Natl. Acad. Sci* 92, 11105-11109.
- Moore, A. W., Schedl, A., McInnes, L., Doyle, M., Hecksher-Sorensen, J., Hastie, N. D. (1998). YAC transgenic analysis reveals *Wilms' Tumour 1* gene activity in the proliferating coelomic epithelium, developing diaphragm and limb. *Mechanisms of Development* 79, 169-184.
- Morita, R., Saito, S., Ishikawa, J., Ogawa, O., Yoshida, O., Yamakawa, K., and Nakamura, Y. (1991). Common regions of deletion on chromosomes 5q, 6q, and 10q in renal cell carcinoma. *Cancer Research* 51, 5817-5820.
- Morris, J. F., Madden, S. L., Tournay, O. E., Cook, D. M., Sukhatme, V. P., and Rauscher III, F. J. (1991). Characterization of the zinc finger protein encoded by the WT1 Wilms' tumor locus. *Oncogene* 6, 2339-2348.
- Moshier, J. A., Skunca, M., Wu, W., Boppana, S. M., Rauscher III, F. J., and Dosescu, J. (1996). Regulation of ornithine decarboxylase gene expression by the Wilms' tumor suppressor WT1. *Nucleic Acids Research* 24, 1149-1157.
- Nachtigal, M. W., Hirokawa, Y., Enyeart-VanHouten, D. L., Flanagan, J. N., Hammer, G. D., and Ingraham, H. A. (1998). Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. *Cell* 93, 445-454.
- Nagase, T., Miyajima, N., Tanaka, A., Sazuka, T., Seki, N., Sato, S., Tabata, S., Ishikawa, K., Kawarabayasi, Y., Kotani, H., and Nomura, N. (1995). Prediction of the coding sequences of unidentified human genes. III. The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Research* 2, 37-43.
- Ogawa, O., Eccles, M. R., Szeto, J., McNoe, L. A., Yun, K., Maw, M. A., Smith, P. J., and Reeve, A. E. (1993). Relaxation of Insulin-Like Growth Factor-I Gene Imprinting Implicated in Wilms-Tumor. *Nature* 362, 749-751.
- Ohlsson, R., Nystrom, A., Pfeifer-Ohlsson, S., Tohonen, V., Hedborg, F., Schofield, P., Flam, F., and Ekstrom, T. J. (1993). IGF2 is parentally imprinted during human embryogenesis and in the Beckwith-Wiedemann syndrome. *Nature Genetics* 4, 94-97.
- Overall, M. L., Spencer, J., Bakker, M., Dziadek, M., and Smith, P. J. (1996). p57K1P2 is expressed in Wilms' tumor with LOH of 11p15.5. *Genes Chromosomes Cancer* 17, 56-59.
- Park, S., Schalling, M., Bernard, A., Maheswaran, S., Shipley, G. C., Roberts, D., Fletcher, J., Shipman, R., Rheinwald, J., Demetri, G., and et al (1993). The Wilms tumour gene WT1 is expressed in murine mesoderm-derived tissues and mutated in a human mesothelioma. *Nature Genetics* 4, 415-420.
- Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809-817.
- Pelletier, J., Bruening, W., Kashtan, C. E., Mauer, S. M., Manivel, J. C., Striegel, J. E., Houghton, D. C., Junien, C., Habib, R., Fouser, L., and et al (1991a). Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. *Cell* 67, 437-447.
- Pelletier, J., Schalling, M., Buckler, A. J., Rogers, A., Haber, D. A., and Housman, D. (1991b). Expression of the Wilms' tumor gene WT1 in the murine urogenital system. *Genes & Development* 5, 1345-1356.
- Perlman, M., Goldberg, G. M., Bar-Ziv, J., and Danovitch, G. (1973). Renal haemartomas and nephroblastomas with fetal gigantism: a familial syndrome. *Journal Pediatrics* 83, 414-418.

Pichel, J. G., Shen, L., Sheng, H. Z., Granholm, A.-C., Drago, J., Grinberg, A., Lee, E. J., Huang, S. P., Saarma, M., Hoffer, B. J., Sariola, H., and Westphal, H. (1996). Defects in enteric innervation and kidney development in mice lacking GDNF. *Nature* 382, 73-76.

Pritchard Jones, K., Fleming, S., Davidson, D., Bickmore, W., Porteous, D., Gosden, C., Bard, J., Buckler, A., Pelletier, J., Housman, D., van Heyningen, V., and Hastie, N. (1990). The candidate Wilms' tumour gene is involved in genitourinary development. *Nature* 346, 194-197.

Pritchard Jones, K., and Hastie, N. D. (1990). Wilms' tumour as a paradigm for the relationship of cancer to development. *Cancer Survival* 9, 555-578.

Rahman, N., Arbour, L., Tonin, P., Baruchel, S., Pritchard-Jones, K., Narod, S. A., and Stratton, M. R. (1997). The familial Wilms' tumour susceptibility gene, FWT1, may not be a tumour suppressor gene. *Oncogene* 14, 3099-3102.

Rahman, N., Arbour, L., Tonin, P., Renshaw, J., Pelletier, J., Baruchel, S., Pritchard-Jones, K., Stratton, M., and Narod, S. A. (1996). Evidence for a familial Wilms' tumour gene (FWT1) on chromosome 17q12-q21. *Nature Genetics* 13, 461-463.

Ramkisson, Y., and Goodfellow, P. (1996). Early steps in mammalian sex determination. *Current Opinion in Genetics & Development* 6, 316-321.

Rauscher, F. J., Morris, J. F., Tournay, O. E., Cook, D. M., and Curran, T. (1990). Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. *Science* 250, 1259-1262.

Reddy, J. C., Hosono, S., and Licht, J. D. (1995a). The transcriptional effect of Wt1 is modulated by choice of expression vector. *Journal of Biological Chemistry* 270, 29976-29982.

Reddy, J. C., Morris, J. C., Wang, J., English, M. A., Haber, D. A., Shi, Y., and Licht, J. D. (1995b). WT1-mediated transcriptional activation is inhibited by dominant negative mutant proteins. *The Journal of Biological Chemistry* 270, 10878-10884.

Reeve, A. E., Eccles, M. R., Wilkins, R. J., Bell, G. I., and Millow, L. J. (1985). Expression of insulin growth factor-II transcripts in Wilms' tumor. *Nature* 317, 258-260.

Reeve, A. E., Sih, S. A., Raizis, A. M., and Feinberg, A. P. (1989). Loss of allelic heterozygosity at a second locus on chromosome 11 in sporadic Wilms' tumor cells. *Molecular and Cellular Biology* 9, 1799-1803.

Riccardi, V. M., Sujansky, E., Smith, A. C., and Franke, U. (1978). Chromosomal imbalance in the aniridia-Wilms' tumour association: 11p interstitial deletion. *Pediatrics* 61, 604-610.

Rio-Tsonis, K. D., Covarrubias, L., Kent, J., Hastie, N., and Tsonis, P. A. (1996). Regulation of the Wilms' tumor gene during spermatogenesis. *Developmental Dynamics* 207, 372-381.

Rose, E. A., Glaser, T., Jones, C., Smith, C. L., Lewis, W. H., Call, K. M., Minden, M., Champagne, E., Bonetta, L., Yeger, H., and al, e. (1990). Complete physical map of the WAGR region of 11p13 localises a candidate Wilms' tumor gene. *Cell* 60, 495-508.

Rupprecht, H. D., Drummond, I. A., Madden, S. L., Rauscher, F. J. I., and Sukhatme, V. P. (1994). The Wilms' tumor suppressor gene, WT1, is negatively autoregulated. *The Journal of Biological Chemistry* 269 (8), 6198-6206.

Ryan, G., Steeleperkins, V., Morris, J. F., Rauscher, F. J., and Dressler, G. R. (1995). Repression of Pax-2 By Wt1 During Normal Kidney Development. *Development* 121, 867-875.

Saxen, L. (1987). *Organogenesis of the Kidney* (Cambridge U.K.: Cambridge University Press).

Schedl, A., and Hastie, N. (1998). Multiple roles for the Wilms' tumour suppressor gene, WT1 in genitourinary development. *Molecular and Cellular Endocrinology* 140, 65-69.

- Scholz, H., Bossone, S. A., Cohen, H. T., Akella, U., Strauss, W. M., and Sukhatme, V. P. (1997). A far upstream cis-element is required for Wilms' Tumour-1 (WT1) gene expression in renal cell culture. *The Journal of Biological Chemistry* 272, 32836-32846.
- Schroeder, W. T., Chao, L. Y., Dao, D. D., Strong, L. C., Pathak, S., Riccardi, V., Lewis, W. H., and Saunders, G. F. (1987). Non random loss of maternal chromosome-11 alleles in Wilms-Tumors. *American Journal of Human Genetics* 40, 413-420.
- Sells, S. F., Han, S. S., Muthukkumar, S., Maddiwar, N., Johnstone, R., Boghaert, E., Gillis, D., Liu, G., Nair, P., Monnig, S., Collini, P., Mattson, M. P., Sukhatme, V. P., Zimmer, S. G., Wood Jr, D. P., McRoberts, J. W., Shi, Y., and Rangnekar, V. M. (1997). Expression and function of the leucine zipper protein Par-4 in apoptosis. *Molecular and Cellular Biology* 17, 3823-3832.
- Seufert, W., Futcher, B., and Jentsch, S. (1995). Role of ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature* 373, 78-81.
- Sharma, P. M., Bowman, M., Madden, S. L., Rauscher, F. J., and Sukumar, S. L. (1994). RNA editing in the Wilms' tumor susceptibility gene, WT1. *Genes and Development* 8, 720-731.
- Spector, D. L. (1996). Nuclear organization and gene expression. *Experimental Cell Research* 229, 189-197.
- Spector, D. L., Fu, X. D., and Maniatis, T. (1991). Associations between distinct pre-messenger-RNA splicing components and the cell-nucleus. *EMBO Journal* 10, 3467-3481.
- The European Backcross Collaborative Group. (1994) Towards high resolution maps of the mouse and human genomes - a facility for ordering markers to 0.1cM resolution. *Human Molecular Genetics* 3 621-627.
- Turner, B. M., and Franchi, L. (1987). Identification of protein antigens associated with the nuclear matrix and with clusters of interchromatin granules in both interphase and mitotic cells. *Journal of Cell Science* 87, 269-282.
- Wang, Z. Y., Madden, S. L., Deuel, T. F., and Rauscher, F. J. (1992). The Wilms' tumor gene product, WT1, represses transcription of the platelet-derived growth factor A-chain gene. *The Journal of Biological Chemistry* 267, 21999-22002.
- Wang, Z. Y., Qiu, Q. Q., and Deuel, T. F. (1993). The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains. *The Journal of Biological Chemistry* 268, 9172-9175.
- Wang, Z.-Y., Qiu, Q.-Q., Seufert, W., Taguche, T., Testa, J. R., Whitmore, S. A., Callen, D. F., Welsh, D., Shenk, T., and Deuel, T. F. (1996). Molecular cloning of the cDNA and chromosome localization of the gene for human ubiquitin-conjugating enzyme 9. *The Journal of Biological Chemistry* 271, 24811-24816.
- Weiner, J. S., Coppes, M. J., and Ritchey, M. L. (1998). Current concepts in the biology and management of Wilms' tumour. *The Journal of Urology* 159, 1316-1325.
- Werner, H., Re, G. G., Drummond, I. A., Sukhatme, V. P., Rauscher, F. J., Sens, D. A., Garvin, A. J., LeRoith, D., and Roberts, C. T. J. (1993). Increased expression of the insulin-like growth factor I receptor gene, IGF1R, in Wilms tumor is correlated with modulation of IGF1R promoter activity by the WT1 Wilms tumor gene product. *Proc.Natl.Acad.Sci.U.S.A.* 90, 5828-5832.
- Wu, Y. J., Frazier, G. C., and Saunders, G. F. (1995). GATA-1 transactivates the WT1 hematopoietic specific enhancer. *The Journal of Biological Chemistry* 270, 5944-5949.
- Zamore, P. D., Patton, J. G., and Green, M. R. (1992). Cloning and domain-structure of the mammalian splicing factor U2AF. *Nature* 355, 609-614.

Appendix

Appendix: Primers used in this thesis

The following table contains all the primers used in the thesis, and a brief description of why they were used.

Primer	Sequence 5' -3'	Purpose
H792	AGA GAT CTC TCG AGG ATC	5' primer for pGAD10 vector.
H793	CAG TAT CTA CGA TTC ATA	3' primer for pGAD10 vector
I163	TGG CTT ACC CAT ACG ATG	Sequencing pAS vector
J782	TCC CCC GGG TCA GTA TCT ACG ATT CAT AGA	3' primer adding extra restriction sites to clone insert from pGAD10 into pAS.
J783	AAC CAT GCC ATG GAG ATC TCT CGA GGA TCC	5' primer adding extra restriction sites to clone insert from pGAD10 into pAS.
L55	AGC TGT TGG AGG ACT TAT TAC	Primer starting at position 941bp to sequence NAL in reverse direction.
L56	AGA GAT GCA AGA GTG TAC TAC	Primer starting at position 248bp of NAL to sequence the clone further.
L342	CTG AGT GGC TGG AAG TTT AC	Primer starting at position 416bp of NAL to sequence the clone further
L343	CTG GCG TGT CTC CTT CAG CTG	Primer starting at position 694bp of NAL to sequence the clone in the reverse direction
L344	TTC TAG GAA GGA TAG AGT GTA	Primer used to sequence clone B
L345	TTC CAA GGA CCT CCA GGA TAT	Primer used to further sequence clone B
L497	CGT ACC TCA GGA AGC AGG TTT	3' primer used to sequence clone A
L957	AAC CAC TGT CACCTG GTT GGA	pGAD10 primer upstream of H792, to check the frame of clones.
M715	CGG AAT TCT CAT TGG CTA TCA GGC GTA AAC TT	3' primer adding EcoRI site and stop codons in frame at position 428bp to subclone the N-terminal half of NAL back into pGAD10.
M716	GGA ATT CAC AGG GAA AAA GTT ATT AAT GGC GAA G	5' primer adding EcoRI site in frame at position 450bp of NAL to subclone the C-terminus back into pGAD10.
M717	CCC AAG AAG GTT CGA TTG AGT AAA CAG ACT	5' primer starting at position 14bp on NAL used for RT-PCR

Primer	Sequence 5' -3'	Purpose
M718	ACT TCT GAA CGT GGC GGG AAC CCA CAG TTC	3' primer starting at position 1142bp on NAL used for RT-PCR
M728	AAG TAA TAT AGA GCC CCA AAC TAA ATG GTT	3' primer starting at position 1271bp of EST D14661 used for RT-PCR
M729	TGCCCA ACT GAG ATC AAC AAT GGT AGA CCC	5' primer starting at position 306bp of NAL used for RT-PCR
M730	TAC CAA GCT GGA CAG TGT AAG AAT GTG GCT	3' primer starting at position 724bp of the EST D14661 used for RT-PCR
M731	TCC TGA CTG CTT TTA AGC TCC TCA CTG TAT	3' primer starting at position 597bp of NAL used for RT-PCR
N842	GCT CTA GAG AAG AAC CTC TTC CC	5' primer used to subclone NAL out of pGAD10 and into pCIneo- HA
N845	ACG CGT CGA CCT AGC CTC GAC TGT CA T	3' primer used to subclone NAL out of pGAD10 and into pCIneo- HA
N966	ATT AAC CCT CAC TAA AGG GA	T3 primer to sequence the pCIneo- HA-NAL vector
N967	TAA TAC GAC TCA CTA TAG G	T7 primer to sequence the pCIneo- HA-NAL vector
P475	TAA TGG CGA AGT GTC GAA TGC	5' primer used to sequence Nal from 418bp
P725	AAA GAC TGC AGT CGT CTG	5' primer to sequence Nal from 751bp
Q13	GTT CAG AAT GGC TTG GAC TCA	5' primer starting at position 1132bp on NAL
Q147	TCA GTG CGG GGT ATG AAA GTG	5' primer starting at position 987bp on NAL
Q148	ATT CAG ACT TGT TAG AGT AGA	5' primer starting at position 1497bp on Nal
R385	GTC AAA AAA TTA ATG CAT AC	5' primer starting at position 1194bp of Nal, used for EUCIB mapping
R386	CAG TAC AGG TGA GAT ATA C	3' primer starting at position 1743bp of Nal used for EUCIB mapping